

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:18:11 ; Search time 2502.29 Seconds  
(without alignments)  
8935.871 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgaacatgtcccaaaaca.....gtagttataaaaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estm:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	120	13.0	895	CNS0071A	AL066286 Drosophil
C 2	113	12.3	712	13	BX416727
C 3	111.4	12.1	1180	13	BX436369
C 4	104.6	11.4	946	29	AG127412 Pan trogl

C 5	103.2	11.2	922	29	CNS0073W
C 6	101.8	11.1	919	29	CNS04ENY
C 7	100	10.9	1201	9	AL581589
C 8	99.8	10.8	1201	9	AL531977
C 9	99.4	10.8	1141	28	AQ743305
C 10	99	10.8	1101	29	CNS00FXE
C 11	98	10.7	1151	14	CD501049
C 12	97.2	10.6	1200	9	AL548181
C 13	96.6	10.5	1024	29	AG133080
C 14	96.2	10.5	753	13	BX443342
C 15	96.2	10.5	1100	13	BX385076
C 16	96.2	10.5	1172	29	BZ696157
C 17	95.8	10.4	1097	29	CNS01224
C 18	95.8	10.4	1101	29	CNS0100X
C 19	95	10.3	791	29	CNS009KS
C 20	95	10.3	1101	29	CNS00LQO
C 21	95	10.3	1144	28	AQ743364
C 22	94.6	10.3	899	13	BX436853
C 23	94.4	10.3	1216	29	AG135357
C 24	94.2	10.2	815	28	B12686
C 25	94	10.2	1000	13	BX418086
C 26	94	10.2	1188	29	AG135332
C 27	93.8	10.2	859	29	AG128925
C 28	93.8	10.2	1008	29	AG137085
C 29	93.8	10.2	1099	28	AQ743360
C 30	93.6	10.2	884	29	CNS008UO
C 31	93.6	10.2	969	28	AQ743309
C 32	93.6	10.2	1081	29	AG135328
C 33	92.8	10.1	955	29	AG076494
C 34	92.8	10.1	981	29	AG127518
C 35	92.8	10.1	1183	29	BZ696114
C 36	92.8	10.1	1353	29	AG128010
C 37	92.6	10.1	923	13	BX410248
C 38	92.6	10.1	1183	29	AG136828
C 39	92.4	10.0	880	29	AG139490
C 40	92.2	10.0	885	29	CNS031TM
C 41	92	10.0	1101	29	CNS001T2
C 42	92	10.0	1137	28	AQ743326
C 43	92	10.0	1225	29	CNS0166K
C 44	92	10.0	1491	10	BE882936
C 45	91.4	9.9	700	29	AG127423

#### ALIGNMENTS

RESULT 1	CNS0071A/c	895 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BAC14B09 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL066286				
VERSION	AL066286.1	GI:4945153			
KEYWORDS	GSS				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 895)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of				



AUTHORS	Li,W.B., Gruber,C.; Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 516.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/custer.cgi?seq=CSOCAP001BG06QP1&cluster=516.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Paraday Avenue Genome science ID : CSOCAP001BG06QP1.  Location/Qualifiers 1 . 1180 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOCAP001YM12" /tissue type="THYMUSt" /clone_lib="Homo sapiens THYMUS" /note="Vector: pCMVSPORT 6; lsc strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT	127 a 154 C 191 G 420 t 288 others
ORIGIN	
Query Match	12.1%; Score 111.4; DB 13; Length 1180;
Best Local Similarity	25.8%; Pred. No. 1e-08;
Matches	144; Conservative 194; Mismatches 219; Indels 1; Gaps 1;
QY	325 AATTACATCAAAACCAGCACAATACTACTGTTTCACAAAACACACGGAGTGTAAGTCAAACCT 384
Db	1152 RAAAAAAGAANAIAAAAAAAAAARVAAMAAASAMAAMMAAAAMAAAVAAAMAAAMMAAAAA 1093
QY	385 GCACCCCAACAGGTCAAGCATTAABAAAACACACACACACCAACCAACCAACCGCACGCC 444
Db	1092 AAAAIAAAAAAGMAMAAAGMAMMMMAAAAAAAAAAAAAAAAAAAAAVMMAAAAAAMMAAAA 1033
QY	445 CACTAACAAACACGCACGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTCGA 504
Db	1032 AGAANWNAWAAAAANWAAMMAAMMANWMAWMMAWWNNNMMAAMMACMWAKAAAAA 973
QY	505 AGTGTTTAATCTTGATCCCTGACGATATGACGACAACTCAACCTGCTGGGGTATCTG 564
Db	972 AAAAAAAMAAAAAHKHAAMAAAAAMAAAWMAHAMMAAMMAAMMAAMMAAMMAAMNEM 913
QY	565 CAAAGAAATACCAACAAACAAACAGGACGAAAGAAACCAACCAACCAAGCTCAAAAAAAC 624
Db	912 AMMNWANWNABAACMCMAAYMYMMMAWCNCMMMMMMMCMNMCMMAAMMAAAAA 853
QY	625 AACCTTCAAGACACCAAAAAAGATCTCAAACCTCAAACCACTAAACCAAGGAGTAGC 684
Db	852 AAAAAAAAAAAAAAAAAACHBACMMWMAWHMHMMMAHMMAAMMAAMMYMYMMC 793
QY	685 CACCAACAGGCCACAGAGAGCCACCATCAACACCAACCAACCAACCAACCAACCACTAC 744
Db	792 MMWAKYCMMWMAWCHHWAYAAYAAAAAAMMAAAAYAAAAAAMMAAMMBEACYMMNMW 733
QY	745 ACTGCTCACCACAA-CACACACAGGAATCCAACCTCAACAGTCAAAATGGAACCTTCC 803
Db	732 MNMCBAAMBAMMACSACCBSHCMSCKNHAKHHHHHMHCMCYHHHYHBYY 673
QY	804 ACTCAACTCTCTCCGAGGCACTCTAGCCCCCTTCTCAAGTCTCCACACACATCCGAGCAC 863
Db	672 YCYWBYBFCHKICTAAAAAAMMAAAAKYYYYYAAAAANRKYNYMYMKYTMAANY 613
QY	864 CATCAACACCTCTCATCTC 881

Db	612 AAAAAAGHYNTXKHVTYY 595
RESULT 4	
AG127412	
LOCUS	AG127412 946 bp DNA linear GSS 04-NOV-2001
DEFINITION	Pan troglodytes DNA, clone: PTB-138E06.R, genomic survey sequence.
ACCESSION	AG127412
VERSION	AG127412.1 GI:16656577
KEYWORDS	GSS.
SOURCE	Pan troglodytes (chimpanzee)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
AUTHORS	1 BAC end sequences of Library PTB Unpublished
TITLE	2 (bases 1 to 946)
JOURNAL	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
REFERENCE	Direct Submission Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgpc.gsc.riken.go.jp/, Tel:+81-45-503-9111, Fax:81-45-503-9170]
AUTHORS	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
TITLE	PRIMERS
JOURNAL	Sequencing: MJ3Rev
FEATURES	source location/qualifiers
vector	: pkx145
R.Site 1	: SacI
R.Site 2	: SacI
/organism="Pan troglodytes"	
/mol_type="genomic DNA"	
/db_xref="taxon:9598"	
/clone="PTB-138E06.R"	
/sex="male"	
/cell_type="lymphoblast"	
/clone lib="PTB Chimpanzee Male BAC Library"	
BASE COUNT	505 a 351 c 31 g 23 t 36 others
ORIGIN	
Query Match	11.4%; Score 104.6; DB 29; Length 946;
Best Local Similarity	46.6%; Pred. No. 1.3e-07;
Matches	326; Conservative 0; Mismatches 374; Indels 0; Gaps 0;
QY	220 AAACATGTCATCATACAAGATGCACAGCCGAGATCAGACACACCCCACATACCT 279
Db	153 AACAAAAACACACCAACCAACCAACCAACCAACCAANNAAAAACCCACACCAAC 212
QY	280 CACTCAGATCCTCAGCTTGGAATCAGTCTTCCAATCTGTCTGTAATTACATCAACAC 339
Db	213 AGCAACCAAAAAACCAAAAAACAAACAAACAAACAAACAAACAAACAAACAAAC 272
QY	340 CACCACATCTAGCTTCAACAAACACAGAGTCAAGTCAAACTGCAACCCCAACAGT 399
Db	273 CACCCCAACAAAAAACCNAAACAAACCAACCAACCAACCAACCAACCAACAA 332
QY	400 CAAGCTAAAAACACACACACACCAACCAACCAACCAACCAACCAACCAACCAAC 459
Db	333 AAAAATCTATACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 392
QY	460 CCAAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 519
Db	393 AAAACACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 452

```

QY 520 ACCCTGAGCATATGACGACAAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAA 579
Db 453 AAAACAAACCATCTCTCCCAAAACAAACAAACAAACAAACAAACAAACAAACAAAC 512
QY 580 CAAAAAACAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 639
Db 513 AAAAAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 572
QY 640 CAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACT 699
Db 573 GAAAAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 632
QY 700 AGAAGAGCCACCACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 759
Db 633 AAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 692
QY 760 CACCACAGGAATCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACT 819
Db 693 AAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752
QY 820 AGGCAATCTAAGCCCTTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGT 879
Db 753 CAGCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 812
QY 880 TCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 919
Db 813 CCAAAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 852

```

RESULT 5  
CNS0073W/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BAC14D09 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 922)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osogawa and  
Aaron Mammeter in Peter de Jong's laboratory in the Department of  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
1..922  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BAC14D09"  
/note="end : TET3"

BASE COUNT 223 a 95 c 109 g 221 t 274 others

## ORIGIN

Query Match 11.2%; Score 103.2; DB 29; Length 922;  
Best Local Similarity 24.8%; Pred. No. 2.3e-07;  
Matches 102; Conservative 164; Mismatches 144; Indels 1; Gaps 1;  
QY 386 CAACCCACACATCTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCAACCA 445  
Db 921 MMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 862  
QY 446 ACTTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505  
Db 861 ACANWAAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 802  
QY 506 GTGTTTAACTTTGTACCTGTCAGCATATGTCAGCAACCAATCCCAACCTGCTGGGTATCTGC 565  
Db 801 ACACWCAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 743  
QY 566 AAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
Db 742 AWAAMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 683  
QY 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 685  
Db 682 MWCAWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 623  
QY 686 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 745  
Db 622 AMWACWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 563  
QY 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 796  
Db 562 AMWACWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 512

RESULT 6  
CNS04ENY/c

LOCUS  
DEFINITION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)



- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

**COMMENT**

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8262.f. For more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=C90DG004CC08NP1&cluster=8262.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=C90DG004CC08NP1&cluster=8262.f). Contact :  
Feng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
Faraday Avenue Genoscope sequence ID : C90DG004CC08NP1

## FEATURES

**URES**  
**source**

```

BASE COUNT
99 a      86 c      254 g      409 t      71 others
/move- genome sequence ID : COBG104DH0/LPI~end : T/

```

[illegible]

## RESULT 7

AL581589/c	AL581589	Homo sapiens	1201 bp	mrna	linear	EST 01-JUN-2003
LOCUS	AL581589	Homo sapiens B CELLS (RAMOS CELL LINE)				
DEFINITION	clone CSODG004YF15 3-PRIME, mrna sequence.					
ACCESSION	AL581589					
VERSION	AL581589.1	GI:12948741				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1201)					
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.					
TITLE	Full-length cDNA libraries and normalization					
JOURNAL	Unpublished					

**Dib**



[illegible]









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 00:01:13 ; Search time 310.043 Seconds  
(without alignments)  
8070.528 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacatgtccaaaca.....gtagttattaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1811591 seqs, 1359896290 residues

Total number of hits satisfying chosen parameters: 3623182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	853.4	92.8	15223	10	US-09-847-173-1
2	637.4	69.3	696	10	US-09-844-645-4
3	441.4	48.0	15225	12	US-09-827-688-10
C 4	86.2	9.4	7758	12	US-10-311-455-1076
C 5	85.8	9.3	793	10	US-09-878-574-4304
C 6	85.8	9.3	3673778	12	US-10-312-841-2
C 7	84.2	9.2	1372	12	US-10-017-161-1981
C 8	81.4	8.8	16167	12	US-10-311-455-1056
C 9	81.4	8.8	16167	12	US-10-240-485-82
C 10	81	8.8	6270	12	US-10-311-455-378
C 11	81	8.8	6270	12	US-10-240-453-20
C 12	81	8.8	6270	14	US-10-239-676-12
C 13	80	8.7	6668	12	US-10-311-455-1670
C 14	78.2	8.5	1635	9	US-09-864-761-20241
C 15	78.2	8.5	1973	9	US-09-864-761-3471
C 16	78.2	8.5	2543	12	US-10-029-386-20536

C 17	76.6	8.3	576	10	US-09-878-574-4296	Sequence 4296, Ap
C 18	76.4	8.3	529	10	US-09-983-963-2109	Sequence 2109, Ap
C 19	76.4	8.3	1030	10	US-09-878-574-4306	Sequence 4306, Ap
C 20	76.4	8.3	1075	9	US-09-864-761-19241	Sequence 19241, A
C 21	76.4	8.3	1403	9	US-09-864-761-2513	Sequence 2513, Ap
C 22	75.2	8.2	6668	12	US-10-311-455-1669	Sequence 1669, Ap
C 23	75	8.2	577	10	US-09-878-574-4312	Sequence 4312, Ap
C 24	74.8	8.1	6089	12	US-10-311-455-675	Sequence 675, App
C 25	74.8	8.1	11622	12	US-10-311-455-648	Sequence 648, App
C 26	74.8	8.1	14006	12	US-10-311-455-1931	Sequence 1931, App
C 27	74	8.0	639	10	US-09-878-574-4316	Sequence 4316, Ap
C 28	74	8.0	1493	12	US-10-029-386-25133	Sequence 25133, A
C 29	73.6	8.0	1083	12	US-10-029-386-20723	Sequence 20723, A
C 30	73.6	8.0	16170	12	US-10-311-455-1241	Sequence 1241, Ap
C 31	72.8	7.9	6048	12	US-10-311-455-2004	Sequence 2004, Ap
C 32	72.2	7.8	13606	12	US-10-311-455-1783	Sequence 1783, Ap
C 33	72.2	7.8	13606	12	US-10-240-453-187	Sequence 187, App
C 34	72.2	7.8	13606	12	US-10-239-676-165	Sequence 165, App
C 35	71.8	7.8	3163	14	US-10-017-161-1857	Sequence 1857, App
C 36	71.2	7.7	3673778	12	US-10-312-841-1	Sequence 1, Appli
C 37	71	7.7	13732	12	US-10-311-455-1793	Sequence 1793, Ap
C 38	70.8	7.7	11996	12	US-10-240-485-45	Sequence 45, Appl
C 39	70.6	7.7	8147	12	US-10-311-455-401	Sequence 401, App
C 40	70.6	7.7	9415	12	US-10-311-455-268	Sequence 268, App
C 41	70.4	7.7	8149	12	US-10-311-455-1123	Sequence 1123, Ap
C 42	70.4	7.7	9539	12	US-10-240-453-54	Sequence 54, Appl
C 43	70.4	7.7	9539	14	US-10-239-676-52	Sequence 52, Appl
C 44	70.2	7.6	10467	12	US-10-240-453-327	Sequence 327, App
C 45	70.2	7.6	17869	12	US-10-311-455-78	Sequence 78, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-847-173-1  
; Sequence 1, Application US/09847173  
; Publication No. US20020182228A1  
; GENERAL INFORMATION:  
; APPLICANT: Collins, Peter L.  
; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
; SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/847,173  
; FILING DATE: 03-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/720,132  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 15280-250-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-467-9600  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15223 base pairs  
; TYPE: nucleic acid



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

Query Match      92.8%; Score 853.4; DB 10; Length 15223;
Best Local Similarity 95.5%; Pred. No. 1.8e-210;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAAAACATGTCGCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 60
DB 4682 TGCAAAACATGTCGCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGA 4741
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 120
DB 4742 CACTCTCAATCATTTATTTATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 4801
QY 121 AGCACAATACATATTCATCTGCGCAATGATTAATCTCACTTCACTTATATTAATTCAGC 180
DB 4802 AGCACAATACATATTCATCTGCGCAATGATTAATCTCACTTCACTTATATTAATTCAGC 4861
QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGGTACACTCACTCACTTCACTTATATTAATTCAGC 240
DB 4862 CATCATATTCATAGCTCGGCAACCAACCAAGGTACACTCACTCACTTCACTTATATTAATTCAGC 4921
QY 241 TGCACCAAGCGAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300
DB 4922 TGCACCAAGCGAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 4981
QY 301 AATCAGCTTCTCAATCTGCTGCAATTAATCATACCAACCAACCACTTCACTTCACTTCAAC 360
DB 4982 AATCAGCTTCTCAATCTGCTGCAATTAATCATACCAACCAACCACTTCACTTCACTTCAAC 5041
QY 361 AACACCAAGGAGTCAAGTCAAACTTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
DB 5042 AACACCAAGGAGTCAAGTCAAACTTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5101
QY 421 AACCAAAACCAACCCAGAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 480
DB 5102 AACCAAAACCAACCCAGAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5161
QY 481 ACCCAATATGATTTTCACTTCGAGTGTTCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
DB 5162 ACCCAATATGATTTTCACTTCGAGTGTTCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5221
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 600
DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 5281
QY 601 CACCAACCAAGCTTCAAAACCAACCACTTCAAGACCAACCAAAAGATCTCAAACTCA 660
DB 5282 CACTACCAAGGCCCAACCAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 5341
QY 661 AACCACTTAAACCAAGGAGTACCCACCAACCAAGCCCAAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 5342 AACCACTTAAACCAAGGAGTACCCACCAACCAAGCCCAAGGAGGAGGAGGAGGAGGAGGAGG 5401
QY 721 CACCAAAACCAACCAATCACTACCTGCTCAACCAACCAACCAACCAAGAAATCCAAACT 780
DB 5402 CACCAAAACCAACCAATCACTACCTGCTCAACCAACCAACCAAGAAATCCAAACT 5461
QY 781 CAAAGTCAATGAAACCTTTCCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
DB 5462 CAAAGTCAATGAAACCTTTCCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 5521
QY 841 AGTCTCCACACATCCGAGGAGCCATCAACACCTTCTTCCACCCACCAACCAAGCCCA 900
DB 5522 AGTCTCTACACATCCGAGGAGCCATCAACACCTTCTTCCACCCACCAACCAAGCCCA 900
QY 901 GTAGTTATTAAAAA 919
DB 5582 GTAGTTATTAAAAACATA 5600
```

```

RESULT 2
US-09-844-645-4
; Sequence 4, Application US/09844645
; Patent No. US20020102242A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
;              McDaniel, Larry S.
;              Curiel, David T.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
;              ADMINISTERING PNEUMOCOCCAL DNA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/844,645
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-844-645-4

Query Match      69.3%; Score 637.4; DB 10; Length 696;
Best Local Similarity 96.1%; Pred. No. 3.7e-155;
Matches 684; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 8 ATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGACACTCTC 67
DB 1 ATGTCCAAAACCAAGACCAACGACCGCTAAGACACTAGAAAGACCTGGGACACTCTC 60
QY 68 AATCATTTATTTATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTTAGCAAA 127
DB 61 AATCATTTATTTATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTTAGCAAA 120
QY 128 ATCACAATTTATTCATCTGCGCAATGATAATCTCACTTCACTTATATTAATTCAGCCATCATA 187
DB 121 ATCACAATTTATTCATCTGCGCAATGATAATCTCACTTCACTTATATTAATTCAGCCATCATA 180
QY 188 TTCAATAGCTCGGCAAAACCAACCAAGTCACTAAACCACTGCAATCATACAGATGCAACA 247
DB 181 TTCAATAGCTCGGCAAAACCAACCAAGTCACTAAACCACTGCAATCATACAGATGCAACA 240
QY 248 AGCCAGATCAAGAACCAACCAACCCCAACATACCTCACTCAGGATCTTCAAGTCTGGAATCAGC 307
DB 241 AGCCAGATCAAGAACCAACCAACCCCAACATACCTCACTCAGGATCTTCAAGTCTGGAATCAGT 300
QY 308 TTCTCCAAATCTGCTGAAATTAATCATCAAAACCAACCACTACTAGCTTCAACACACCA 367
```

Db 301 CCCTCTAATCCGCTGAAATTTACATCAAAATCACCACCTACTAGCTTCAACACACCA 360  
QY 368 GGAGTCAAGTCAAACTGCAACCCACCAACAGTCAAGACTAAACACCAAC-ACCCCA 426  
Db 361 GGAGTCAAGTCAAACTGCAACCCACCAACAGTCAAGACTAAACACCAACAACTCA 420  
QY 427 AACACAAACCCAGCAAGCCCACTCAAAACCAACAGCCCAAAACCAACCAACCCAA 486  
Db 421 AACACAAACCCAGCAAGCCCACTCAAAACCAACAGCCCAAAACCAACCAACCCAA 480  
QY 487 TAATGATTTTCACTTCAAGTGTAACTTTGTACCTGAGCATATGCAAGCAAAATCC 546  
Db 481 TAATGATTTTCACTTCAAGTGTAACTTTGTACCTGAGCATATGCAAGCAAAATCC 540  
QY 547 AACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGCAAGCAAGCAAGCCAC 606  
Db 541 AACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGCAAGCAAGCAAGCCAC 600  
QY 607 CAAGCCTTACAAAAACCAACCTTCAAGACACCAAAACCAAAAGATCTCAAACTCAAAACCCAC 666  
Db 601 CAAGCCTTACAAAAACCAACCTTCAAGACACCAAAACCAAAAGATCTCAAACTCAAAACCCAC 660  
QY 667 TAAACCAAGCAAGTACCCACCAAGCC 697  
Db 661 TAAATCAAGCAAGTACCCACCAAGCC 691

RESULT 3  
US-09-827-688-10  
; Sequence 10, Application US/09827688  
; Publication No. US20030165476A1  
; GENERAL INFORMATION:  
; APPLICANT: ORSON, FRANK  
; APPLICANT: KINSEY, BERNA  
; APPLICANT: BHOGAL, BALIR  
; TITLE OF INVENTION: MACROORGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION  
; FILE REFERENCE: P01949US1/1004014  
; CURRENT APPLICATION NUMBER: US/09/827,688  
; CURRENT FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/195,680  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 15225  
; TYPE: DNA  
; ORGANISM: RSV  
US-09-827-688-10

Query Match 48.0%; Score 441.4; DB 12; Length 15225;  
Best Local Similarity 68.1%; Pred. No. 1.1e-103;  
Matches 629; Conservative 0; Mismatches 291; Indels 3; Gaps 1;

QY 1 TCCAAACATGTCCAAAAACAGGACCAACGCGCTAAGACACTAGAAAAAGCTGGGA 60  
Db 4683 TCCAAACATGTCCAAAAACAGGACCAACGCGCTAAGACACTAGAAAAAGCTGGGA 4742  
QY 61 CACTCTCAATCATTTATTTATTCATATCATATCGGGCTTATATAGTTAAATCTTAT 120  
Db 4743 TACTCTCAATCATTTATTTATTCATATCATATCGGGCTTATATAGTTAAATCTTAT 4802  
QY 121 AGCACAATATCAATATTCATATTCGCAATGATATCTCAATCACTTATATTAATAC 180  
Db 4803 AGCACAATATCAATATTCGCAATGATATCTCAATCACTTATATTAATAC 4862  
QY 181 CATCATATTCATATTCGCAATGATATCTCAATCACTTATATTAATAC 240  
Db 4863 CATATATTCATATTCGCAATGATATCTCAATCACTTATATTAATAC 4922  
QY 241 TCCAAACAGCCAGATCAAGAAACCAACCCCACTCACTCAGGATCTTCAGCTGG 300

Db 4923 AATAAAAAACCACTGAAAAAATACATCACCACCTACTTACTCAAGTCCCAACAGAAAG 4982  
QY 301 AATCAGCTTCTCAATCTGTAATTTACATCAAAACCAACCAATAGCTTCAAC 360  
Db 4983 GGTTAGTCTATCCAAACCAACCTTACCAATCCACCAATTCACCAATTCAGCCCAAC 5042  
QY 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACCAAC 420  
Db 5043 ATCAACCAACCAAGTCAAGAAACACCAACCAACCAACCAACCAAGGCAAGAACCC 5102  
QY 421 AACCCAAACCAACCCAGCAAGCCCACTCAAAACCAACGCAAAACCAACCAACCAAC 480  
Db 5103 CAGCTCAACACAGACCAACCAAGCCGAGCAAAACCCAGCTTAAACCAACCAACCA 5162  
QY 481 ACCATAATGATTTTCACTTCAAGTGTAACTTTGTACCTGAGCATATGCAAGCA 540  
Db 5163 ACCAAAGATGATTTTCACTTCAAGTGTAACTTTGTACCTGAGCATATGCAAGCA 5222  
QY 541 CAATCAACCTGCTGGGCTATCTGCAAAAGATACCAACCAAAACCAACCAAGCAAGAAAC 600  
Db 5223 CAATCACTTTGCAATTCATCTGTAACAAATACCAACCAACCAACCAAGCAAGAAAC 5282  
QY 601 CACCAACAGCTTCAAAAAACCAACCTTCAA---GACACCAAAAAAGATCTCAAAAC 657  
Db 5283 AACCATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5342  
QY 658 TCBAACCACTAAACCAAGGAAGTACCCACCAACCAAGCCCAACAGAGCCCAACCAAC 717  
Db 5343 ACCAGCCAAACGAGCAAAAAAGAACTACCAACCAACCAACCAACCAACCAACCAAC 5402  
QY 718 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 777  
Db 5403 GACCAAGAAAGACACCAAGCACTTCAATTCCTGCTGCAACCAACCAACCAACCA 5462  
QY 778 ACTCAAGTCAATGGAACCTTCACTCAACCTCTCCGAGGCAATCTTAAGCCCTTC 837  
Db 5463 ACACCAATTCACAGCAATCTCTCACTCAACCCCGGAAACCAACCAACCAACCAAC 5522  
QY 838 TCAAGTCTCCCAACATCGGAGCAGCCATCAACACCTCATCTCCACCCCAACCAACAG 897  
Db 5523 ACAAACCCCAAGCATCCGAGCCCTTACATCAATTCACCAAAATTCACCAATTCACA 5582  
QY 898 CCAGTAGTTATTAAAAAAGAA 920  
Db 5583 TGCTTAGTTATCAAAACTACA 5605

RESULT 4  
US-10-311-455-1076/c  
; Sequence 1076, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1076  
; LENGTH: 7758  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1076

Query Match  
Best Local Similarity 9.4%; Score 86.2; DB 12; Length 7758;  
Matches 312; Conservative 0; Mismatches 338; Indels 3; Gaps 2;  
QY 121 AGCACAATTCATATTCATTCGCAATGATATCTCAATCTCACTTTATATTAACAGC 180  
DB 1866 AACAAAAAACAACCAAACTCCCAAAACACAAATATAAACAATAAACHCTTAAC 1807  
QY 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTTAACAACTGCAATCATACAAGA 240  
DB 1806 AACAAAAAACAACCTTAACATCTTAACACAAACAAAAAACAACCTTAACAACTA 1747  
QY 241 TGCACACGCGAGTCAAGACACAAACCCCACTCACTCACTCACTCACTCACTCACT 300  
DB 1746 AACAAAAAACAACCTTAACACCTTAACACCAAAAAAACAACCTTAACACCTTA 1687  
QY 301 AATCAGCTTCTCAATCTGCTGAAATTTACATCAACAAACCCACCATAGCTTCAAC 360  
DB 1686 ACCAAAAAACAACCTTAACCTCCCAAAACACAAATTAATTAATTAACCTTA 1627  
QY 361 AACACCGAGGTCAGTCAAGTCAACCTCGCAACCCCAACAGTCAAGTCAAGTCAAG 420  
DB 1626 AACCAACAAAAACAACCAAAACACCTTAACAAAAAACAACCTTAACCACTTA 1567  
QY 421 AACCAACCAACCAACCCAGCCCACTTACAAACACGCGCAAAACCAACCAACCA 480  
DB 1566 AACCAACAAAAACAACCTTAACCACTTAACCAACAAAAAACAACCTTAACCA 1507  
QY 481 ACCCAAT--AATGATTTTCACTTCGAAGTGTGTTAACTTTGACCTGCGAGCATATG 538  
DB 1506 AACAAATTAACCTTAACCAACCAAAATTAACCACTTAACCACTTAACCAAA 1447  
QY 539 AACAAATCAACCTGCTGGCTATCTGC--AAAAGAAATACCAACAAAAAACCAGGA 597  
DB 1446 AACAACTAAACATCTAAACCAACCAACAAACAAATTAACAACTTAACCAAA 1387  
QY 598 AACCAACCAACCAACCAACCAACCAACCTTCAAGACACCAACCAAAAGATCTCA 657  
DB 1386 AACCAACCAACCAACCAACCAACCAACCAACCTTAACCAACCAACCAACCA 1327  
QY 658 TCAAAACCACTTAACCAAGGAGTACCCACCAACCAACCCACAGAGCCCAACCATCA 717  
DB 1326 ACCAATTAACCACTTAACCAACCAACCAACCAACCACTTAACCAACCAACCA 1267  
QY 718 CACCAACCAACCAACCAACCAACCAACCTGCTCAACCAACCAACCAACCAAG 770  
DB 1266 AACCAACCAACCAACCTTAACCAACCAACCAACCAACCAACCAACCAACCA 1214

RESULT 5  
US-09-878-574-4304/c  
; Sequence 4304, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4304  
; LENGTH: 793  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

; LOCATION: (1)..(793)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-C11  
US-09-878-574-4304

Query Match  
Best Local Similarity 9.3%; Score 85.8; DB 10; Length 793;  
Matches 283; Conservative 0; Mismatches 300; Indels 2; Gaps 2;  
QY 191 ATAGCTCTGGCAAAACCAAAAGTCACTAACTGCAATCATATCAAGATGCAACAGC 250  
DB 715 ACAAAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 656  
QY 251 CAGATCAAGACACAAACCCCA-ACATACCTCACTCAGATCTCTCAGCTTGGATCAGTT 309  
DB 655 CAACACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 596  
QY 310 CTCCAATCTGTCTGAAATTTACATCAACAAAC-CACCACCATACTAGTTTCAACACAC 368  
DB 595 ACAAAACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 536  
QY 369 GAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 428  
DB 535 NAACAAACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 476  
QY 429 CACAACCCGCAAGCCCACTTACAAACCAACGCGCAAAACCAACCAACCAACCAAC 488  
DB 475 CAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 416  
QY 489 ATGATTTTCACTTCTGAAAGTGTGTTAACTTTGACCTGCGAGCATATGCGAACAAT 548  
DB 415 ACAACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 356  
QY 549 CCTGCTGGCTTATCTGCAAAAGAAATACCAACAAAAAACCAGGAAGAAACCAACCA 608  
DB 355 ACACACACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 296  
QY 609 AGCTTCAAAAAAACCACCTTCAAGACCAACCAACCAACCAACCAACCAACCAAC 668  
DB 295 CACCAAAACAAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 236  
QY 669 AACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 728  
DB 235 ACACACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 176  
QY 729 CAACATCAACAACTACCTGCTCAACCAACCAACCAACCAACCAACCAACCAACCA 773  
DB 175 CACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 131

RESULT 6  
US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2  
Query Match  
9.3%; Score 85.8; DB 12; Length 3673778;

```
Best Local Similarity 47.0%; Pred. No. 2.8e-10;
Matches 334; Conservative 0; Mismatches 372; Indels 5; Gaps 2;

QY 214 CACACTAACCACTGCAATCATACAGATGCAAGCCAGATCAAGCAACCAACCCCAAC 273
Db 2435804 CTCACAAACCTTACAAACACGAAACCAAAATCTCCACTAACCAATAAAACCAACCAACC 2435745

QY 274 ATACCTCACTCAGGATCTCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAAATTAATC 333
Db 2435744 ATCCCGACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435685

QY 334 ACAACACCAACCACTAGCTTCAACCAACACGAGGAGTCAAGTCAAACTGCAACCCAC 393
Db 2435684 CCAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435625

QY 394 AACAGTCAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 453
Db 2435624 AACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435565

QY 454 ACAACGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 513
Db 2435564 AAAC-CTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435506

QY 514 CTTTGTACCTGAGCATGAGCAACATCAACCTGCTGGGCTATCTGCAAAAGAT 573
Db 2435505 CAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435446

QY 574 ACCAAACAAA---AACAGGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 629
Db 2435445 CCCAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435386

QY 630 TCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 689
Db 2435385 CAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435326

QY 690 CCAAGCCCAAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTGC 749
Db 2435325 CACAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435266

QY 750 TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTTCACTCA 809
Db 2435265 TAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435206

QY 810 CTTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCGAGCACCATCAC 869
Db 2435205 CCTCAACAAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435146

QY 870 AACCTCATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
Db 2435145 CATCCCAACCAACCAACCTTACAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA 2435095
```

## RESULT 7

```
US-10-017-161-1981
; Sequence 1981, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1981
; LENGTH: 1972
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1972)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1772)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (975)..(394)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1319)..(1328)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-1981
```

```
Query Match 9.2%; Score 84.2; DB 12; Length 1972;
Best Local Similarity 45.4%; Pred. No. 1.6e-11;
Matches 350; Conservative 0; Mismatches 418; Indels 3; Gaps 3;
```

```
QY 128 ATCAATATCCATTTCTGGCAATGATAATCTCAACTTCTACTTATATATTACAGCCATCATTA 187
Db 594 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 653

QY 188 TTCAATAGCTCGGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 247
Db 654 ATCAACGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 713

QY 248 AGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 307
Db 714 ATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 773

QY 308 TTCTCCAATCTGTGAAATTAATCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 366
Db 774 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 833

QY 367 AGGAGTCAAGTCAAAACCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 425
Db 834 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 893

QY 426 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 485
Db 894 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 953

QY 486 ATAAATGATTTTCACTTCGAGAGTTTAACTTTGTTAGCTTGGAGCATATGAGCAACCAATC 545
Db 954 ACCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1013

QY 546 CAACCTGTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAAACCAAAACCAAAACCAACCA 605
Db 1014 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1073

QY 606 CCAAGCCTACAAAACCAACCAACCTTCAAGACCAACCAAAACCAAAACCAAAACCAAAACCAACCA 665
Db 1074 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1133

QY 666 CTAACCAAGGAAGTATCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 725
Db 1134 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1193

QY 726 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 785
Db 1194 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1253

QY 786 GTCAATGGAAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGGCTTCTCAAGTCT 845
Db 1254 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1312

QY 846 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 896
Db 1313 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1363
```

```
RESULT 8
US-10-311-455-1056/c
; Sequence 1056, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1056
; LENGTH: 16167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732
; OTHER INFORMATION: n is a or g or c or t
; NAME/KEY: unsure
; LOCATION: 5837..5838
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1056

Query Match      8.8%; Score 81.4; DB 12; Length 16167;
Best Local Similarity 46.6%; Pred. No. 2.5e-10;
Matches 235; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

Qy 314 AATCTGTCTGAAATTACATCAAAACCAACCACTACTAGTTTCAACCAACCAAGGAGTC 373
Db 5866 ATTCTTACTTTTCTCACACCAAAACGNNNAACAAACGAAAAAAGAAAAAAGAAAAA 5807
Qy 374 AAGTCAAACCTGCAACCCCAACAGCTCAAGCTTAAACACACACACACACCCCAACCA 433
Db 5806 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5747
Qy 434 CCCAGCAAGCCCTACTACAAACCAACGCAAAACCAACCAACCAACCAACCAATATGAT 493
Db 5746 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5687
Qy 494 TTTCACTTCGAAAGTTTAACTTTGTACCTGCGCATATGCGACCAATCCAACTGCG 553
Db 5686 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5627
Qy 554 TGGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAGGCAACCAACCAACCAAC 613
Db 5566 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5507
Qy 674 AAGGAAGTACCAACCAACCAACCAAGAGAGGCAACCAATCAACCAACCAACCAAC 733
Db 5506 AACCTTTTACCAACCAACCAATTAATCATCATCACCCTTAACCTTAAACCTTAA 5447
Qy 734 ATCACTACTACTGCTCAACCAACCAACCAAGGAGGCAACCAATCAACCAACCAAC 793
Db 5446 TCTCCCCCAATTCCTCTCAACCTTTCTTTATCAACCAACCTTAAATATAAATAA 5387
Qy 794 GAAACCTTCCACTCAACCTCTCTCC 817
```

```
Db 5386 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5363

RESULT 9
US-10-240-485-82/c
; Sequence 82, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240.485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 82
; LENGTH: 16167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732)
; NAME/KEY: unsure
; LOCATION: (5837..)
US-10-240-485-82

Query Match      8.8%; Score 81.4; DB 12; Length 16167;
Best Local Similarity 46.6%; Pred. No. 2.5e-10;
Matches 235; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 314 AATCTGTCTGAAATTACATCAAAACCAACCACTACTAGTTTCAACCAACCAAGGAGTC 373
Db 5866 ATTCTTACTTTTCTCACACCAAAACGNNNAACAAACGAAAAAAGAAAAAAGAAAAA 5807
Qy 374 AAGTCAAACCTGCAACCCCAACAGCTCAAGCTTAAACACACACACACCCCAACCA 433
Db 5806 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5747
Qy 434 CCCAGCAAGCCCTACTACAAACCAACGCAAAACCAACCAACCAACCAACCAATATGAT 493
Db 5746 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5687
Qy 494 TTTCACTTCGAAAGTTTAACTTTGTACCTGCGCATATGCGACCAATCCAACTGCG 553
Db 5686 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5627
Qy 554 TGGGCTATCTGCAAGAAATACCAACCAAAACCAAGGAGGCAACCAACCAACCAAC 613
Db 5626 CGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5567
Qy 614 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 673
Db 5566 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 5507
Qy 674 AAGGAAGTACCAACCAACCAACCAAGAGAGGCAACCAATCAACCAACCAACCAAC 733
Db 5506 AACCTTTTACCAACCAACCAATTAATCATCATCACCCTTAACCTTAAACCTTAA 5447
Qy 734 ATCACTACTACTGCTCAACCAACCAACCAAGGAGGCAACCAATCAACCAACCAAC 793
Db 5446 TCTCCCCCAATTCCTCTCAACCTTTCTTTATCAACCAACCTTAAATATAAATAA 5387
Qy 794 GAAACCTTCCACTCAACCTCTCTCC 817
```

Qy 734 ATCACTACTACCTGCTCTCAACCAACACCCAGGAAATCTCAAACCTCACAAGTCAATG 793  
 |||||  
 Db 5446 TCTCCCCCAATTCCTCAAACTCTTATCAACCCCTAATAAATAAAAAAAAAAAAAA 5387  
 |||||  
 Qy 794 GAAACCTTCCCACTCAACCTCTCTCC 817  
 |||||  
 Db 5386 AAAAAAAAAAAAAATAAATCTCTCTCC 5363  
 |||||

```

RESULT 10
US-10-311-455-378/c
; Sequence 378, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 378
; LENGTH: 6270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-378

```

Query Match	8.8%	Score 81	DB 12	Length 6270
Best Local Similarity	45.1%	Pred. No. 2e-10		
Matches 345	Conservative 0	Mismatches 415	Indels 5	Gaps 1
QY	38	AAGACACTAGAAAGACCTGGGACACTCTCAATCATTTATTATTTCATATCATCGGGCTTA	97	
Db	2056	AACATACTCTTAAACACCCGCAAACTCGGGTATATAAAACCTATTCTCTTAAACCTATCACT	1997	
QY	98	TATAAGTTAAATCTTAAATCTGTAGCACAATCATATTCCTTCTGGCAATCATATTC	157	
Db	1996	ACTCCCTTAAACCTCTAAAAACCAACCCCTTAACCTACCTCAAAAAATATCTCTAAAAAAT	1937	
QY	158	TCAACTTCACCTTATAATTACAGCCATCATATTCATAGCCTCGGCAACCAACCAAGTCACA	217	
Db	1936	AAAACTTACAAAAAAACCTTTCTCTTAAAAATAATAAAACCCACACAAAAAAACCA	1877	
QY	218	CTAACAACTGCAATCATAGAAGTGCACAAAGCCAGATCAAGAACACAAACCCCAACATC	277	
Db	1876	AAAAAAACCGTATCAAAAAAATAAAACCCCAACTAAAAATAAAACACTAAAAAAAC	1817	
QY	278	CTCACTCAGGATCCTCAGCTTTGGAATCAGCTTCTCC-----AATCTGTCTGAATTTACAT	332	
Db	1816	TTTCTCTCTCACTCTTAACCTTAAACCTCACTTCTACATAAAACTTTTTCAAAATCGAAA	1757	
QY	333	CACAAACCAACCAACCATCTAGCTTTCAACACACCCAGGAGTCAAGTCAAACTGCAACCCA	392	
Db	1756	AAAAAAATTACTCAAATAAATAATATTACCTCTACGAAATTTTAAATCCCTTAACCTCC	1697	
QY	393	CAACAGTCAAGACTAAAAACAAACAAACCAACCAACCAACCAAGCCCACTACAA	452	
Db	1696	CAACAAAAATAAAAAAACAATCCCAAAAAACATCACTTCGAAACATCTCCAAACCCCTAA	1637	
QY	453	AACAAGCCAAAAACAACCAACCAACCAACCAACCAATATGATTTTTCATCTCGAAGTGTTA	512	
Db	1636	CTTAATCCCTTAACCAAAAAATATCCCAACTACTTAATTAACCAAAATCGCCTTCTCT	1577	

QY	513	ACTTTGTACCTCGAGCATATGCGAGCAACAAATCAACCTGTGTGGCTATCTGCAAAAGAA	572
Db	1576	AAAACTTACATCTCTCCCTCTAAATATATCTACTCCCTCCTCTAAATCAACCCCAAAAAA	1517
QY	573	TACCAACAAAAACAGGAGAAAGAAACCCACCAACCAAGCTTACAAAAAAACCAACCTTCA	632
Db	1516	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATCTACAAAAAATAAAAAAAAAAACA	1457
QY	633	AGACACCAAAAAAGATCTCAAACTCAAACTCAAAACCAAGGAAGTAGTACCACCAACCA	692
Db	1456	AAACAAAAAATAAAAAAAAAAACAACACAAAAATAATTAACAAAAAACTACAAAAACAA	1397
QY	693	AGCCACAGAGAGCCAACCATCAACACCCCAACCAACACATCACAACCTACACTGTCTCA	752
Db	1396	ATCTTAAACAAAAAATCGAAAAAATAACAAACCCCAATCAAAAAATAAAAAAAAAAACA	1337
QY	753	CCAAACAACACACAGGAAATCCAAAACTCAAGTCAAAATGGAAA	797
Db	1336	ACCGAAACACCCCAAAAAATACCTCAAAAAACAATAAATAAACGAAA	1292

```

RESULT 11
US-10-240-453-20/c
; Sequence 20, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPNEROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation St
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240.453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 20
; LENGTH: 6270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapi
US-10-240-453-20

```

	Query Match	8.8%	Score 81;	DB 12;	Length 6270;
	Best Local Similarity	45.1%;	Prod. No. 2e-10;		
	Matches 345;	Conservative 0;	Mismatches 415;	Indels 5;	Gaps 1;
QY	38	AAGACACTAGAAAGACCTGGGACACTCTCAATCATTATTATTCATATCATATCGGGCTTA	97		
Db	2056	AACATACTCTAAACACACGCGAACTTCGGGTATAAACTTATTCCTTAAACACTATCACT	1997		
QY	98	TATAAGTTAACTTTAAATCTGTAGCACAAATCACATTATTCCTACTTCGGCAATGATAATC	157		
Db	1996	ACTCCCTAAACCCCTCAAAAAACAAACCCCTAACTACACTCAAAAAATCTCTAAATAAT	1937		
QY	158	TCAACTTTCATTATAATTACAGCCATCATATTCATAGCTTCGCGAACCAACCAAGTCACA	217		
Db	1936	AAAACTTCAAAAAAAACTTTCTCTTAAAAATAATAATAAAAACCCCAACCAAAAAAACCA	1877		
QY	218	CTAAACACTGCATCATACAAGATGCAACAAGCCAGATCAAGAACCAACACCCCAACAATAC	277		

```
Db 1876 AAAAAAGCTATCAAAAAAATAAATTAACCACTAAATAATTAACCTAAAAAAGC 1817
QY 278 CTCACCTCAGAGCTCTCAGCTTGGATTCAGCTTCTCC-----AATCTGTCTGAAATTTACAT 332
Db 1816 TTTCTCTCTTCACTCTAACTTAAACCTCAATTTCTACATATAAACTTTTTCATAAATCGAAA 1757
QY 333 CACAAACCAACCACTACTAGCTTCAACCAACCAAGAGTCAAGTCAAACTGCGAACCA 392
Db 1756 AAAAAAATTTACTCAAAATAAATAATTTTACCTCTACGAAAAATTTAAATCCCTAACCTCC 1697
QY 393 CAACAGTCAAGTCAAAAAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 452
Db 1696 CAACAAAAATTAATAAATAATTTCCCAAAACATCACTTCGAACATCTCCAAAAACCTTAA 1637
QY 453 AACAAGCGCAAAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 512
Db 1636 CTTAATCCCTAACCAAAAAATATCCCACTACTAAATAAACAACCAAAATCGCTTCTCT 1577
QY 513 ACTTTGTACCTCGCAGCATATGCGACCAACATCCACCTGCTGGGCTATCTGCAAAAGAA 572
Db 1576 AAAAACTTACATCTCTCCCTTAAATATATCTACCCCTCCCTAAATCAACCCCAAAAAA 1517
QY 573 TACCAACAAAAACCAAGCAAGAAACCAACCAACCAAGCTTCAAAAAACCAACCTTCA 632
Db 1516 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1457
QY 633 AGACAACCAAAAAAGATCTCAAACTTCAACCACTAAACCACTAAACCAAGAGTACCCACCA 692
Db 1456 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1397
QY 693 AGCCACAGAGGCGCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752
Db 1396 ATCTAAACAAAAATCGAAAAAATAAACAACCAACCAACCAACCAACCAACCAACCAACCA 1337
QY 753 CCAACAACCAACCAAGAAATCCAAACTCAACAGTCAAAATGGA 797
Db 1336 ACGAACAACCCCAAAATATCTCAAAACCAATAAACAAGAA 1292
```

## RESULT 12

```
US-10-239-676-12/c
; Sequence 12, Application US/10239676
; Publication No. US2003008260A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013,1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 12
; LENGTH: 6270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-12
```

Query Match 8.8%; Score 81; DB 14; Length 6270;  
Best Local Similarity 45.1%; Pred. No. 2e-10;

```
Matches 345; Conservative 0; Mismatches 415; Indels 5; Gaps 1;
QY 38 AAGACACTAGAAAAAGACCTGGGACACTCTCAATCAATTTATTATTATTCATATATCGGGCTTTA 97
Db 2056 AACATACCTTAAACACCAACCAACCTGCGGTATATAAACTTTATTCCTAAACCTATCACT 1997
QY 98 TATAAGTTTAAATCTTAAATCTGTAGCACAATATCAATTTATCGATTTCTGGCAATGATATC 157
Db 1996 ACTCCCTAAACCCCTCAAAAAACCAACCCCTAACCTACACTCAAAAAATATCTCTAAAAAT 1937
QY 158 TCAACTTCACTTATATTTACAGCCATCATATTTCTATAGCTCGGCAACCAACCAAGTCA 217
Db 1936 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1877
QY 218 CTAACAATGCAATCATACAAGATGCAACCAAGCCAGATCAAGAACACCAACCCCAACATAC 277
Db 1876 AAAAAAAGCTATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1817
QY 278 CTCACCTCAGGATCTCTCAGCTTGGATCAGCTTCTCC-----AATCTGTCTGAAATTTACAT 332
Db 1816 TTTCTCTCTTCACTCTAACTTAAACCCCTCAATTTCTACATAAAAACTTTTTCAAAAATCGAAA 1757
QY 333 CACAAACCAACCAACCATACTAGCTTCAACAACCAACCAAGGAGTCAAGTCAAAACCTGCAACCCA 392
Db 1756 AAAAAAATTTACTCAAAATAAATAATTTACCTCTACGAAAAATTTAAATCCCTTAACCTCC 1697
QY 393 CAACAGTCAAGACTAAAAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 452
Db 1696 CAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1637
QY 453 AACACGCGCAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 512
Db 1636 CTTAATCCCTAACCAAAAAATATCCCACTACTAAATAAACAACCAAAATCGCTTCTCT 1577
QY 513 ACTTTGTACCTCGCAGCATATGCGACCAACATCCAACTCTGCGCTATCTGCAAAAGAA 572
Db 1576 AAAAACTTACATCTCTCCCTTAAATAATATATCTACCCCTCCCTAAATCAACCCCAAAAAA 1517
QY 573 TACCAACAAAAACCAAGCAAGAAACCAACCAACCAAGCTTCAAAAAACCAACCTTCA 632
Db 1516 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1457
QY 633 AGACAACCAAAAAAGATCTCAAACTTCAACCACTAAACCACTAAACCAAGAGTACCCACCA 692
Db 1456 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1397
QY 693 AGCCACAGAGGCGCAACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752
Db 1396 ATCTAAACAAAAATCGAAAAAATAAACAACCAACCAACCAACCAACCAACCAACCAACCA 1337
QY 753 CCAACAACCAACCAAGAAATCCAAACTCAACAGTCAAAATGGA 797
Db 1336 ACGAACAACCCCAAAATATCTCAAAACCAATAAACAAGAA 1292
```

## RESULT 13

```
US-10-311-455-1670/c
; Sequence 1670, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine
; FILE REFERENCE: 5013,1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
```



```
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1670
; LENGTH: 6668
; TYPE: DNA
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 1936
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1670

Query Match
Best Local Similarity 8.7%; Score 80; DB 12; Length 6668;
Matches 215; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 358 AACACACACGAGGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTAAACACACAC 417
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 AACACCCAAACACACCCGAGCCGCTCAACAAACAAACGCCAAACAAACCCACAA 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 CAACCCCAATGATTTTCACTTCGAGGTGTTTAACTTTGTACCTGCGCATATGCG 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CAACCAATCAACCTGCTGGCTATCTGCAAAAGATACCAACCAAAACCCAGGAAGAA 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 AACCAACCCAGCTTACAAACAAACCAACCTTCAAGCAACCAACCAAGATCTCAACC 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 TCAACCACTAACCAAGGAGTACCCACCAACCCAGCCACAGAGCCACCATCA 717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 CACCAACCAACCAACATCACTACACTGCTCACCACCAACCAACCAAGGATCCAAA 777
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 ACTCAAGTCAATGGAAA 797
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2931 AAAACCAAAACCAACCAAA 2912
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-864-761-20241
; Sequence 20241, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
```

```
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20241
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006547.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-20241

Query Match
Best Local Similarity 8.5%; Score 78.2; DB 9; Length 1635;
Matches 354; Conservative 0; Mismatches 398; Indels 7; Gaps 3;

QY 128 ATCACTTATTCATTCTGGCAATGATTAATCTCAACTTTCACCTTAATTAATACAGCCATCAT 187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 ACCACCAATACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCA 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 TTCTAGCTCTGGCAACCAACCAAGTCACTAACTCACTCACTCACTCACTCACTCACTCACT 247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 437 ACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCA 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 AGCCAGATCAAGACACACACCCCACTACCTCACTCAGGATCTCTAGTTGGATCAGC 307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 497 ATCACTCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCA 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 TTCTCCAATCTGTCTGAAATTTACATCAAAACCAACCACTACTAGCTTCAACCAACCA 367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 557 ATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCACTACCATCACT 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 GGAGTCAAGTCAAACTCTGCAACCCCAACCAAGTCAAGCTAAAGCAACCAACCAACCA 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 617 ACCACCACTCACTCTGCGCACCCACCACTCACTCACTCACTCACTCACTCACTCACTCACT 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 428 ACACACCCAGCAGCCCACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAAT 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

Db 677 ACTGCGACACACACACACACACAT-CATTCACACACACACACATCACCACACACAC 735
Qy 488 AATGATTTTCACTTGGAGTGTTTAACTTTGTACCTGCGAGATATGCGACACAAATCCA 547
Db 736 CATCACTACACACACAC- - - - -CACCACCATCACTTACCACACACACACACCATCACT 790
Qy 548 ACTGCTGGGTCTCTGAAAGATATACAAACAAAACACAGAAAGAAACACACACAC 607
Db 791 ACCACACACACATCACCACACACACACACATCACTTACCACACACACACACACCATC 850
Qy 608 AAGCTTACAAAACACACCTTCAAGACACACACACACACACACATCTCAACATCTCAACCACT 667
Db 851 ACTACACACACACACACATCACTTACCACACACACACACACATCACTTACCACACACAC 910
Qy 668 AAACCAAGGAAGTACACACACACACACACACACACACACACACACACACACACAC 727
Db 911 ACCACACACATCACTTACCACACACACACACACACATCACTTACCACACACACAC 970
Qy 728 ACAACATCACTACATGCTGACCAAGACACACACACACACACACACACACACACAC 787
Db 971 ACCACATCACTACACACACACACACACATCACTTACCACACACACACACACACAC 1030
Qy 788 CAATGGAACCTTCACTCAACCTCTCCGAGGCAATCTAAGC-CCTTCTCAAGTCTC 846
Db 1031 ACCATCACTACACACACACACACACACACACACACACACATCACTTACCACACACAC 1090
Qy 847 CACACATCGAGCACCACATCAACACCTCTATCTCCACC 885
Db 1091 ACTACACACACACACACACATCACTTACCACACACACACACACAC 1129

```

## RESULT 15

```

US-09-864-761-3471
; Sequence 3471, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3471
; LENGTH: 1973
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006547.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
; US-09-864-761-3471

```

Query Match 8.5%; Score 78.2; DB 9; Length 1973;

Best Local Similarity 46.6%; Pred. No. 5.9e-10;

Matches 354; Conservative 0; Mismatches 398; Indels 7; Gaps 3;

```

Qy 128 ATCAATTATCCATTCTGGCAATGATTAATCTCAACTTCACTTAAATTACAGCCATCATA 187
Db 531 ACCACCATACCATCACTTACCACCATCACTACACACACACACATCACCATCACCACCATC 590
Qy 188 TTGATAGCTGGCAAAACCAAGTCACTAACAATGAACTGGAATCATACAGATGCAACA 247
Db 591 ACCATCACTATACCACACACACACACACCATCACTACACACACACATCACCATCACCACC 650
Qy 248 AGCCAGATCAAGAACACAAACCCCAACATCTCACTCAGGATCTCAGTTGGATCAGC 307
Db 651 ATCACTATCACTACACACACACACACACACACATCACTACACACACACATCACCACCATC 710
Qy 308 TTCTCCAACTGTCTGAAATTTACATCAAAACACACACACATCACTAGTTTCAACACACCA 367
Db 711 ATCAACACACACACACACACACACATCACTTACCACACACACATCACCATCACCCT 770
Qy 368 GGAGTCAAGTCAAACTTGACACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 427
Db 771 ACCACACCATCACTGTCACACACACACACACACACATCACTACACACACACATCACCACCATC 830
Qy 428 ACACACACACACACACACCTTACAAAACACACACACACACACACACACACACACACAC 487
Db 831 ACTGCAACACACACACACACACATCACTACACACACACACACACACACACACACACAC 889
Qy 488 AATGATTTTCACTTCAAGTGTTTTAACTTTGTACCTGCGAGCATATGACAGCAACATCCA 547
Db 890 CATCACTACACACACAC- - - - -CACCACCATCACTTACCACACACACACATCACCCT 944
Qy 548 ACTGCTGGGTATCTGCAAAAGATATACCAAAACACAAACACAGAAAGAAACACACACAC 607
Db 945 ACCACACACACATCACCACACACACACACATCACTACACACACACACACACACACATC 1004
Qy 608 AAGCTTACAAAACACACCTTCAAGACACACACACACACACACATCTCAAACTCTCAACCACT 667
Db 1005 ACTTACACACACACACACATCACTTACCACACACACACACACACATCACTTACACACACAC 1064
Qy 668 AAACCAAGGAAGTATCCACACACACACACACACACACACACACACACACACACACAC 727
Db 1065 ACCACACCATCACTTACCACACACACACACACATCACTTACCACATCACTTACCACATCACT 1124

```

Search completed: October 30, 2003, 03:23:48  
Job time : 319.043 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:37:29 ; Search time 71.4618 Seconds  
(without alignments)  
5682.373 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacacatgtccaaaacaa.....gtagtattataaaaaaaaaa 920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS-COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	920	100.0	920	2	US-08-467-963C-7
2	920	100.0	920	2	US-08-838-189D-7
3	920	100.0	920	3	US-08-852-344D-7
4	920	100.0	920	3	US-08-344-639E-7
5	920	100.0	920	3	US-08-467-969A-7
6	920	100.0	920	3	US-08-467-961A-7
7	920	100.0	920	3	US-08-001-554A-7
8	894	97.2	894	2	US-08-467-963C-28
9	894	97.2	894	2	US-08-838-189D-28
10	894	97.2	894	3	US-08-852-344D-28
11	894	97.2	894	3	US-08-344-639E-28
12	853.4	92.8	15222	2	US-08-801-898A-23
13	853.4	92.8	15222	2	US-08-962-690-12
14	853.4	92.8	15222	2	US-08-892-403A-1
15	853.4	92.8	15223	2	US-08-720-132-1
16	441.4	48.0	15225	2	US-08-892-403A-2
17	212.8	23.1	1050	4	US-08-836-501-76
18	212.8	23.1	1050	4	US-09-628-830-76
19	210.8	22.9	303	3	US-08-721-979A-1
20	210.8	22.9	303	3	US-08-836-504A-1
21	210.8	22.9	303	3	US-08-836-501-1
22	210.8	22.9	303	4	US-09-654-289-1
23	210.8	22.9	303	4	US-09-582-876-1
24	210.8	22.9	303	4	US-09-626-830-1
25	209.6	22.8	1071	3	US-08-836-501-77
26	209.6	22.8	1071	4	US-09-626-830-77
27	207.6	22.6	303	3	US-08-721-979A-3

28 207.6 22.6 303 3 US-08-836-504A-3  
29 207.6 22.6 303 3 US-08-836-501-3  
30 207.6 22.6 303 3 US-09-654-289-3  
31 207.6 22.6 303 4 US-09-582-876-3  
32 207.6 22.6 303 4 US-09-626-830-3  
33 201.2 21.9 303 3 US-08-721-979A-14  
34 201.2 21.9 303 3 US-08-836-501-14  
35 201.2 21.9 303 4 US-09-654-289-14  
36 201.2 21.9 303 4 US-09-582-876-14  
37 201.2 21.9 303 4 US-09-626-830-14  
38 137 14.9 840 5 PCT-US91-08177-12  
39 137 14.9 7323 5 PCT-US91-08177-12  
40 128.6 14.0 183 3 US-08-721-979A-23  
41 128.6 14.0 183 3 US-08-836-501-23  
42 128.6 14.0 183 4 US-09-654-289-23  
43 128.6 14.0 183 4 US-09-582-876-23  
44 128.6 14.0 183 4 US-09-626-830-23  
45 126 13.7 303 3 US-08-721-979A-2

## ALIGNMENTS

### RESULT 1

US-08-467-963C-7

; Sequence 7, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIs:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 920 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-467-963C-7

```
Query Match      100.0%; Score 920; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 8.8e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACCATGTCCTCAAAACCAAGGACCAAGCGCTTAAGACACTAGAAAAAGACTGGGA 60
DB 1 TGCACCATGTCCTCAAAACCAAGGACCAAGCGCTTAAGACACTAGAAAAAGACTGGGA 60

QY 61 CACTCTCAATCAATTTATTATTATATATCATATCATCGGGCTTATATAAGTTAAATCTGT 120
DB 61 CACTCTCAATCAATTTATTATTATATATCATATCATCGGGCTTATATAAGTTAAATCTGT 120

QY 121 AGCACAATTCATATTCATTTCTGGCAATGATATCTCAACTTCACTTAAATATACAGC 180
DB 121 AGCACAATTCATATTCATTTCTGGCAATGATATCTCAACTTCACTTAAATATACAGC 180

QY 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTAACTGCAATCATACAGA 240
DB 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTAACTGCAATCATACAGA 240

QY 241 TGCACCAAGCCAGATCAAGAACCAACCCCAACATCACTCAGGATCTCAGCTGG 300
DB 241 TGCACCAAGCCAGATCAAGAACCAACCCCAACATCACTCAGGATCTCAGCTGG 300

QY 301 AATCAGCTTCTCCAAATCTGTCTGAATTTACATACAAACCCCAACATCACTTCAAC 360
DB 301 AATCAGCTTCTCCAAATCTGTCTGAATTTACATACAAACCCCAACATCACTTCAAC 360

QY 361 AACACGAGGTCAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACCAACAAAC 420
DB 361 AACACGAGGTCAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACCAACAAAC 420

QY 421 AACCCAAAACACACCCAGCAAGCCCACTACAAAACAAAGCCAAAACCAACCAAA 480
DB 421 AACCCAAAACACACCCAGCAAGCCCACTACAAAACAAAGCCAAAACCAACCAAA 480

QY 481 ACCCAATATGATTTTCACTTCGAGTGTGTTAACTTTGTACCTCGCATATGACGAA 540
DB 481 ACCCAATATGATTTTCACTTCGAGTGTGTTAACTTTGTACCTCGCATATGACGAA 540

QY 541 CAATCAACCTGCTGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGGAAAGAAC 600
DB 541 CAATCAACCTGCTGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGGAAAGAAC 600

QY 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGAACCAAAAAAGATCTCAAACTCA 660
DB 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGAACCAAAAAAGATCTCAAACTCA 660

QY 661 AACCACTAAACCAAGGAGTACCCACCAAGCCCAAGCCCAAGAGCCCAACCATCAAC 720
DB 661 AACCACTAAACCAAGGAGTACCCACCAAGCCCAAGCCCAAGAGCCCAACCATCAAC 720

QY 721 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGGAAATCCAAACT 780
DB 721 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGGAAATCCAAACT 780

QY 781 CAAAGTCAAAATGGAACCTTCACTCAACCTCTCCGAAAGGCAATCTAAGCCCTTCTCA 840
DB 781 CAAAGTCAAAATGGAACCTTCACTCAACCTCTCCGAAAGGCAATCTAAGCCCTTCTCA 840

QY 841 AGTCTCAACATTCGAGGACCCCATCAACCTCTCAACCTCTCAACCAACCAAGCCCA 900
DB 841 AGTCTCAACATTCGAGGACCCCATCAACCTCTCAACCTCTCAACCAACCAAGCCCA 900

QY 901 GTAGTTATTAATAAAAAAAA 920
DB 901 GTAGTTATTAATAAAAAAAA 920
```

RESULT 2  
US-08-838-189D-7  
; Sequence 7, Application US/08838189D

```
Patent No. 5998169
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EWASYSHIN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jdb
TELEPHONE: (416) 595-1155
TELEPHONE: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-838-189D-7
```

```
Query Match      100.0%; Score 920; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 8.8e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACCATGTCCTCAAAACCAAGGACCAAGCGCTTAAGACACTAGAAAAAGACTGGGA 60
DB 1 TGCACCATGTCCTCAAAACCAAGGACCAAGCGCTTAAGACACTAGAAAAAGACTGGGA 60

QY 61 CACTCTCAATCAATTTATTATTATATATCATATCATCGGGCTTATATAAGTTAAATCTGT 120
DB 61 CACTCTCAATCAATTTATTATTATATATCATATCATCGGGCTTATATAAGTTAAATCTGT 120

QY 121 AGCACAATTCATATTCATTTCTGGCAATGATATCTCAACTTCACTTAAATATACAGC 180
DB 121 AGCACAATTCATATTCATTTCTGGCAATGATATCTCAACTTCACTTAAATATACAGC 180

QY 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTAACTGCAATCATACAGA 240
DB 181 CATCATATTCATAGCTCGGCAAAACCAAAAGTCACTAACTGCAATCATACAGA 240

QY 241 TGCACCAAGCCAGATCAAGAACCAACCCCAACATCACTCAGGATCTCAGCTGG 300
DB 241 TGCACCAAGCCAGATCAAGAACCAACCCCAACATCACTCAGGATCTCAGCTGG 300

QY 301 AATCAGCTTCTCCAAATCTGTCTGAATTTACATACAAACCCCAACATCACTTCAAC 360
DB 301 AATCAGCTTCTCCAAATCTGTCTGAATTTACATACAAACCCCAACATCACTTCAAC 360
```



QY	781	CACAAGTCAAATGGAAACCTTTCCACTCAACTCTCCGAAGGCAATCTTAAGCCCTTTCTCA	840
Db	781	CACAAGTCAAATGGAAACCTTTCCACTCAACTCTCCGAAGGCAATCTTAAGCCCTTTCTCA	840
QY	841	AGTCTTCCCAACATCCGAGCACCCATCAACAACCTCTATCTCCACCCAAACAACACGCCA	900
Db	841	AGTCTTCCCAACATCCGAGCACCCATCAACAACCTCTATCTCCACCCAAACAACACGCCA	900
QY	901	GTAGTTATTAAAAA	920
Db	901	GTAGTTATTAAAAA	920

## RESULT 4

US-08-344-639E-7  
 / Sequence 7, Application US/08344639E  
 / Patent No. 6033668  
 / GENERAL INFORMATION:  
 / APPLICANT: Klein, Michel H  
 / APPLICANT: Du, Run-Pan  
 / APPLICANT: Ewasyszyn, Mary E  
 / TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
 / TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
 / TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
 / NUMBER OF SEQUENCES: 38  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Sim & McBurney  
 / STREET: 330 University Avenue, 6th Floor  
 / CITY: Toronto  
 / STATE: Ontario  
 / COUNTRY: Canada  
 / ZIP: M5G 1R7

Query Match 100.0%; Score 920; DB 3; Length 920;  
Best Local Similarity 100.0%; Pred. No. 8.8e-243;  
Matches 920; Conservative 0; Mismatches 0; Indels

Db	1	TGCAACATGTC	CGA	AAACAAAGC	AGG	CAACG	CGCTAAG	CACACTA	AGAAAGAC	CTGGGA	60
Qy	61	CACCTCTCAAT	CAATTTATT	TATCATAT	CATCGG	CTTATATA	AGTTAAAT	CTTTAAAT	CTGT	120	
Db	61	CACCTCTCAAT	CAATTTATT	TATCATAT	CATCGG	CTTATATA	AGTTAAAT	CTTTAAAT	CTGT	120	
Qy	121	AGCAAAATC	ACATTTCC	ATTCTGG	CAATGATTAAT	CTCAACTT	CACCTTATA	TAAATACAGC	180		
Db	121	AGCAAAATC	ACATTTCC	ATTCTGG	CAATGATTAAT	CTCAACTT	CACCTTATA	TAAATACAGC	180		
Qy	181	CATCATATT	CTAGCCT	CGGCAAC	CAAAAGT	CCACTCACTA	CAACTGCAAT	CTATACAAGA	240		
Db	181	CATCATATT	CTAGCCT	CGGCAAC	CAAAAGT	CCACTCACTA	CAACTGCAAT	CTATACAAGA	240		
Qy	241	TGCAACAA	GCCAGATCA	AGAACCA	CAACCC	CAACATACCT	CACTCAGGAT	CTCTCAGCTTGG	300		
Db	241	TGCAACAA	GCCAGATCA	AGAACCA	CAACCC	CAACATACCT	CACTCAGGAT	CTCTCAGCTTGG	300		
Qy	301	AATCAGCTT	CTTCCAAT	CTGTCTG	AAATTAAT	CAACAAC	CAACCACTA	CTAGCTTCAAC	360		
Db	301	AATCAGCTT	CTTCCAAT	CTGTCTG	AAATTAAT	CAACAAC	CAACCACTA	CTAGCTTCAAC	360		
Qy	361	AACACAC	GAGGAGTCA	AGTCAAC	CTGCAAC	CCACACAC	AGTCAAGACT	TAATAACACACAC	420		
Db	361	AACACAC	GAGGAGTCA	AGTCAAC	CTGCAAC	CCACACAC	AGTCAAGACT	TAATAACACACAC	420		
Qy	421	AACCCAA	CAACAA	CCCGAG	CGCCACT	TACAAAA	CAACGCAAA	CAACCAACCAACAA	480		
Db	421	AACCCAA	CAACAA	CCCGAG	CGCCACT	TACAAAA	CAACGCAAA	CAACCAACCAACAA	480		
Qy	481	ACCCAA	TAATGATTT	CACTTCG	AAAGTGT	TAACTTTG	TACCTG	CAGCATATG	CAGCAA	540	
Db	481	ACCCAA	TAATGATTT	CACTTCG	AAAGTGT	TAACTTTG	TACCTG	CAGCATATG	CAGCAA	540	
Qy	541	CAATCC	CAACCTG	CTGGCTAT	CTGCAAA	AGATACCA	AAACCAAA	CAAGGAGAAAC	600		
Db	541	CAATCC	CAACCTG	CTGGCTAT	CTGCAAA	AGATACCA	AAACCAAA	CAAGGAGAAAC	600		
Qy	601	CACCA	CAAGCTT	CAAAAAA	CAACCTT	CAAGAC	CAACCAAA	AAAGATCT	CAAACTCA	660	
Db	601	CACCA	CAAGCTT	CAAAAAA	CAACCTT	CAAGAC	CAACCAAA	AAAGATCT	CAAACTCA	660	
Qy	661	AACCACT	AAACCA	AAAGGAGT	TACCC	CAACCA	AGCCCA	CAGAAG	CGCAACCATCA	720	
Db	661	AACCACT	AAACCA	AAAGGAGT	TACCC	CAACCA	AGCCCA	CAGAAG	CGCAACCATCA	720	
Qy	721	CACCA	AAACCA	ACATCA	CACTAC	CTGCTC	ACCAAC	CAACCA	CAGGAAATC	780	
Db	721	CACCA	AAACCA	ACATCA	CACTAC	CTGCTC	ACCAAC	CAACCA	CAGGAAATC	780	
Qy	781	CACAAGT	CAAAATG	AAACCTT	CCACTCA	AACTCT	CTCCGA	AGGCAATCT	TAAAGCCCTTCTCA	840	
Db	781	CACAAGT	CAAAATG	AAACCTT	CCACTCA	AACTCT	CTCCGA	AGGCAATCT	TAAAGCCCTTCTCA	840	
Qy	841	AGTCTCCA	CAACATCC	GAGCAC	CCCATCA	CAACCC	CTCATCT	CCACCG	CAACCAACG	900	
Db	841	AGTCTCCA	CAACATCC	GAGCAC	CCCATCA	CAACCC	CTCATCT	CCACCG	CAACCAACG	900	
Qy	901	GTAGTTAT	TAAAAA	AAAAA	920						
Db	901	GTAGTTAT	TAAAAA	AAAAA	920						

## RESULT 5

```

US-08-467-969A-7
; Sequence 7, Application US/08467969A
; Patent No. 6168786
;
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasystyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21

```

```
/
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: 330 University Avenue, 6th Floor
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/467,969A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/001,554
/ FILING DATE: 06-JAN-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9200117.1
/ FILING DATE: 06-JAN-1992
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-475 MIS:bh
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ TELEX: 065-24567 SIMBAS
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 920 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-467-969A-7

Query Match 100.0%; Score 920; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 8.8e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAAATGTCGAAAACGAGACGACGCGCTAAGACACTAGAAAAGCCTGGGA 60
DB 1 TGCAAAATGTCGAAAACGAGACGACGCGCTAAGACACTAGAAAAGCCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTATATATCATCGGGCTTATATAAGTTAAATCTGT 120
DB 61 CACTCTCAATCATTTATTATTATATATCATCGGGCTTATATAAGTTAAATCTGT 120
QY 121 AGCAAAATCAGATTATTCATTCTGGCAATGATAATCTCAACTTCACTTAATTAACG 180
DB 121 AGCAAAATCAGATTATTCATTCTGGCAATGATAATCTCAACTTCACTTAATTAACG 180
QY 181 CATCATATTCATGCTCGGCAACCAACCAAGTCAGACTAACCTGCAATATACAGA 240
DB 181 CATCATATTCATGCTCGGCAACCAACCAAGTCAGACTAACCTGCAATATACAGA 240
QY 241 TGCAACAGGTCAGATCAAGAACCAACCCCAACATCTCACTCAGGATCCTCAGCTTG 300
DB 241 TGCAACAGGTCAGATCAAGAACCAACCCCAACATCTCACTCAGGATCCTCAGCTTG 300
QY 301 AATCAGCTTCTCCAATCTGTCTGAAATTTACATCACAAACCAACCATCTAGCTTCAAC 360
DB 301 AATCAGCTTCTCCAATCTGTCTGAAATTTACATCACAAACCAACCATCTAGCTTCAAC 360
QY 361 AACACAGGTCAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAG 420
DB 361 AACACAGGTCAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAG 420
```

```
421 AACCCAAACACAAACCCAGCAAGCCCACTACAAAAAACAAGCCAAAAACAACCCAAACAA 480
421 AACCCAAACACAAACCCAGCAAGCCCACTACAAAAAACAAGCCAAAAACAACCCAAACAA 480
481 ACCCAATAATGATTTTTCATCTTGAAGTGTAACTTTGTACCTGTCAGCATATGCAGCAA 540
481 ACCCAATAATGATTTTTCATCTTGAAGTGTAACTTTGTACCTGTCAGCATATGCAGCAA 540
541 CAAATCAAGCTGCTGGCTATCTGCAAAAGATATCCAAACAAAAACAGGAAAGAAAC 600
541 CAAATCAAGCTGCTGGCTATCTGCAAAAGATATCCAAACAAAAACAGGAAAGAAAC 600
601 CACCACCAAGCCTTACAAAAAACAACCTTCAAGACAAACCAAAAAAGATCTCAAACTCA 660
601 CACCACCAAGCCTTACAAAAAACAACCTTCAAGACAAACCAAAAAAGATCTCAAACTCA 660
661 AACCACTAAACCAAGGAAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 720
661 AACCACTAAACCAAGGAAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 720
721 CACCAAAACAAACATCACAACACTACACTGCTCACCACCAACCAAGGAAATCCAAACT 780
721 CACCAAAACAAACATCACAACACTACACTGCTCACCACCAACCAAGGAAATCCAAACT 780
781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
841 AGTCTCCACAAATCCGAGACCCCACTCAACACCTCATCTCCACCAACCAACCAAGCCCA 900
841 AGTCTCCACAAATCCGAGACCCCACTCAACACCTCATCTCCACCAACCAACCAAGCCCA 900
901 GTAGTTATTAAAAAABAAA 920
901 GTAGTTATTAAAAAABAAA 920

RESULT 6
US-08-467-961A-7
; Sequence 7, Application US/08467961A
; Patent No. 6171783
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6TH Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
```

```
; ; REGISTRATION NUMBER: 24,973
; ; REFERENCE/DOCKET NUMBER: 1038-476 MIS:hh
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (416) 595-1155
; ; TELEFAX: (416) 595-1163
; ; TELEX: 065-24567 SIMBAS
; ; INFORMATION FOR SEQ ID NO: 7:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 920 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: DNA (genomic)
; ; US-08-467-961A-7

Query Match 100.0%; Score 920; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 8.8e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACATGTCACAAACAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60
DB 1 TGCACATGTCACAAACAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTATATATATATATATATATATATATATATATATAT 120
DB 61 CACTCTCAATCATTTATTATTATATATATATATATATATATATATATATATATAT 120
QY 121 AGCACAATACATATATATATATATATATATATATATATATATATATATATATAT 180
DB 121 AGCACAATACATATATATATATATATATATATATATATATATATATATATATAT 180
QY 181 CATCATATTTATAGCTTGGGCAACCAAGTCTCACTTAACCTGCAATCATCAAGA 240
DB 181 CATCATATTTATAGCTTGGGCAACCAAGTCTCACTTAACCTGCAATCATCAAGA 240
QY 241 TGCACAGGAGTCAAGTCAAACTTGCACCAAGTCACTCAAGTCACTCAAGTCTGG 300
DB 241 TGCACAGGAGTCAAGTCAAACTTGCACCAAGTCACTCAAGTCACTCAAGTCTGG 300
QY 301 AATCAGTCTTCCAAATCTGTCTGAATTAATCATCAAAACCAACCACTAGCTTCAAC 360
DB 301 AATCAGTCTTCCAAATCTGTCTGAATTAATCATCAAAACCAACCACTAGCTTCAAC 360
QY 361 AACACAGGAGTCAAGTCAAACTTGCACCAAGTCACTCAAGTCACTCAAGTCTGG 420
DB 361 AACACAGGAGTCAAGTCAAACTTGCACCAAGTCACTCAAGTCACTCAAGTCTGG 420
QY 421 AACCCAAACACACACCCAGCAGCCCACTACAAACCAAGCAGCAACCAACCAACAA 480
DB 421 AACCCAAACACACACCCAGCAGCCCACTACAAACCAAGCAGCAACCAACCAACAA 480
QY 481 ACCCAATATATATTTTCACTTGAAGTGTATATTTTGAACCTTGAACCTTGAACCT 540
DB 481 ACCCAATATATATTTTCACTTGAAGTGTATATTTTGAACCTTGAACCTTGAACCT 540
QY 541 CATATCAACTGCTGGGCTATCTGCAAAAGTATCAAAACCAAGTCAAAACCAAGTAA 600
DB 541 CATATCAACTGCTGGGCTATCTGCAAAAGTATCAAAACCAAGTCAAAACCAAGTAA 600
QY 601 CACCAACCAAGCTTACAAAACCAACCTTCAAGACCAACCAACCAACCAACCTTCA 660
DB 601 CACCAACCAAGCTTACAAAACCAACCTTCAAGACCAACCAACCAACCAACCTTCA 660
QY 661 AACCACTAAACCAAGAGTACCCACCAAGGACCAAGGACCAAGGACCAAGGACCAAG 720
DB 661 AACCACTAAACCAAGAGTACCCACCAAGGACCAAGGACCAAGGACCAAGGACCAAG 720
QY 721 CACCAAAACCAACATCACTACCTGCTCAACCAACCAACCAACCAACCAACCAACCT 780
DB 721 CACCAAAACCAACATCACTACCTGCTCAACCAACCAACCAACCAACCAACCAACCT 780
QY 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
DB 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
```

```
DB 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGTCTCCACACATCCGAGCAGCCCATCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCA 900
DB 841 AGTCTCCACACATCCGAGCAGCCCATCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCA 900
QY 901 GTAGTTATTAAAAA 920
DB 901 GTAGTTATTAAAAA 920

RESULT 7
US-08-001-554A-7
; Sequence 7, Application US/08001554A
; Patent No. 6225091
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasahyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JAN-1993
; APPLICATION NUMBER: US/08/001,554A
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-286
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-001-554A-7
```

```
Query Match 100.0%; Score 920; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 8.8e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACATGTCACAAACAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60
DB 1 TGCACATGTCACAAACAGGACCAAGCAGCCGCTAAGACACTAGAAAGACCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTATATATATATATATATATATATATATATATATAT 120
DB 61 CACTCTCAATCATTTATTATTATATATATATATATATATATATATATATATATAT 120
QY 121 AGCACAATACATATATATATATATATATATATATATATATATATATATATATAT 180
DB 121 AGCACAATACATATATATATATATATATATATATATATATATATATATATATAT 180
QY 181 CATCATATTTATAGCTTGGGCAACCAAGTCTCACTTAACCTGCAATCATCAAGA 240
DB 181 CATCATATTTATAGCTTGGGCAACCAAGTCTCACTTAACCTGCAATCATCAAGA 240
```





Db 601 AGCCCTCAAAAACCAACCTTCAAGACACACCAAAAGATCTCAAACTCAAACT 660  
QY 668 AAACCAAGAGAGTACCCACCAACCAAGCCCAAGAGAGAGCCACCAACCAACCA 727  
Db 661 AAACCAAGAGAGTACCCACCAACCAAGCCCAAGAGAGAGCCACCAACCAACCA 720  
QY 728 ACAACATCAACACTACACTGCTCAGCAACCAACCAAGAGAGAGTCAAACTCAAGT 787  
Db 721 ACAACATCAACACTACACTGCTCAGCAACCAACCAAGAGAGTCAAACTCAAGT 780  
QY 788 CAAATGGAACCTTCACTCAACTCAACTCTCGAAGGCAATCTAAGCCCTTCAAGTCTCC 847  
Db 781 CAAATGGAACCTTCACTCAACTCTCGAAGGCAATCTAAGCCCTTCAAGTCTCC 840  
QY 848 ACAACATCGAGACCCCATCAACACCTCATCTCCACCCCAACCAACAGCCAG 901  
Db 841 ACAACATCGAGACCCCATCAACACCTCATCTCCACCCCAACCAACAGCCAG 894

## RESULT 9

US-08-838-189D-28  
; Sequence 28, Application US/08838189D  
; Patent No. 598169  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EWASYSHYN, Mary E  
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,189D  
; FILING DATE: 16-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 894 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-838-189D-28

Query Match 97.2%; Score 894; DB 2; Length 894;  
Best Local Similarity 100.0%; Pred. No. 1.2e-235;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATGTCCTCAAAAACCAAGCCGCAACGCTAAGACACCTAGAAAAGACCTGGGACACTCTC 67  
Db 1 ATGTCCTCAAAAACCAAGCCGCAACGCTAAGACACCTAGAAAAGACCTGGGACACTCTC 60  
QY 68 AATCATTTATTATTATTCATATCATCGGGCTTTATATAAGTTAAATCTTAAATCTGTAGCAAA 127  
Db 61 AATCATTTATTATTATTCATATCATCGGGCTTTATATAAGTTAAATCTTAAATCTGTAGCAAA 120  
QY 128 ATCACTATTATTCATTTCTGCAATGATTAATCTCAACTTCACTTATTAATTAACAGCCATCAT 187  
Db 121 ATCACTATTATTCATTTCTGCAATGATTAATCTCAACTTCACTTATTAATTAACAGCCATCAT 180  
QY 188 TTCTAGTCTGGCAAAACCAACAAAGTCACTAACTAACTGCAATCATATCAAGATGCAACA 247  
Db 181 TTCTAGTCTGGCAAAACCAACAAAGTCACTAACTAACTGCAATCATATCAAGATGCAACA 240  
QY 248 AGCCAGATCAAGAACCAACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGC 307  
Db 241 AGCCAGATCAAGAACCAACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGC 300  
QY 308 TTCTCAATCTCTGTGAAATTAATCAACCAACCAACCACTACTAGTTTCAACCAACCA 367  
Db 301 TTCTCAATCTCTGTGAAATTAATCAACCAACCAACCACTACTAGTTTCAACCAACCA 360  
QY 368 GGAGTCAGTCAAACTGCAACCCCAACAGTCAAGCTTAAATAACACACACCAACCAACCA 427  
Db 361 GGAGTCAGTCAAACTGCAACCCCAACAGTCAAGCTTAAATAACACACCAACCAACCA 420  
QY 428 ACAACACCCAGCAAGCCCACTACAAAACCAACCCCAACCAACCAACCAACCAACCA 487  
Db 421 ACAACACCCAGCAAGCCCACTACAAAACCAACCCCAACCAACCAACCAACCAACCA 480  
QY 489 AATGATTTTCACTTCAAGTGTGTTTAACTTTGTAACCTGCGATATGCAAGCAACCAATCA 547  
Db 481 AATGATTTTCACTTCAAGTGTGTTTAACTTTGTAACCTGCGATATGCAAGCAACCAATCA 540  
QY 548 ACCTGCTGGCTATCTGCAAAAGATATCAACCAACCAACCAACCAACCAACCAACCA 607  
Db 541 ACCTGCTGGCTATCTGCAAAAGATATCAACCAACCAACCAACCAACCAACCAACCA 600  
QY 608 AAGCCTACAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCAAACTCAAAACCACT 667  
Db 601 AAGCCTACAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCAAACTCAAAACCACT 660  
QY 668 AAACCAAGAGAGTACCCACCAACCAAGCCCAAGAGAGCCCAACCATCAACCAACCA 727  
Db 661 AAACCAAGAGAGTACCCACCAACCAAGCCCAAGAGAGCCCAACCATCAACCAACCA 720  
QY 728 ACAACATCAACACTACACTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCT 787  
Db 721 ACAACATCAACACTACACTGCTCACCACCAACCAACCAACCAACCAACCAACCTCAAGT 780  
QY 788 CAAATGGAACCTTCACTCAACTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847  
Db 781 CAAATGGAACCTTCACTCAACTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840  
QY 848 ACAACATCGAGACCCCATCAACACCTCATCTCCACCCCAACCAACAGCCAG 901  
Db 841 ACAACATCGAGACCCCATCAACACCTCATCTCCACCCCAACCAACAGCCAG 894

## RESULT 10

US-08-852-344D-28  
; Sequence 28, Application US/08852344D  
; Patent No. 6017539  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EWASYSHYN, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
; TITLE OF INVENTION: SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 37



; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 894 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-344-639E-28

Query Match 97.2%; Score 894; DB 3; Length 894;

Best Local Similarity 100.0%; Pred. No. 1.2e-235;

Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      8 ATCTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACATCTTC 67
DB      1 ATCTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGACATCTTC 60
QY      68 AATCATTTAATTTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGTAGCACAA 127
DB      61 AATCATTTAATTTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGTAGCACAA 120
QY      128 ATCAGATTATTCATTTCTGGCAATGATAATCTCACTTCACTTATTAATCAGGCATCATA 187
DB      121 ATCAGATTATTCATTTCTGGCAATGATAATCTCACTTCACTTATTAATCAGGCATCATA 180
QY      188 TTCTATAGCTCGGCAACCAACCAAGTCACTACTAACAACCTCAATCTATACAAAGATGCAACA 247
DB      181 TTCTATAGCTCGGCAACCAACCAAGTCACTACTAACAACCTCAATCTATACAAAGATGCAACA 240
QY      248 AGCCAGATCAAGAACCAACCAACCAATACCTCACTCAGGATCCTCAGCTTGGAAATCAGC 307
DB      241 AGCCAGATCAAGAACCAACCAACCAATACCTCACTCAGGATCCTCAGCTTGGAAATCAGC 300
QY      308 TTCTCCATCTGTCTGAATTTAGATCAACCAACCAACCACTACTAGCTTCAACCAACCA 367
DB      301 TTCTCCATCTGTCTGAATTTAGATCAACCAACCAACCACTACTAGCTTCAACCAACCA 360
QY      368 GGAGTCAAGTCAACCTCGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 427
DB      361 GGAGTCAAGTCAACCTCGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
QY      428 ACACACCCAGAGCCCACTACAAAACCAACGCAACCAACCAACCAACCAACCAACCAAC 487
DB      421 ACACACCCAGAGCCCACTACAAAACCAACGCAACCAACCAACCAACCAACCAACCAAC 480
QY      488 AATGATTTTCACTTGGAAAGTGTAACTTTGTACCTTGCAGATATGCGAACAATCCA 547
DB      481 AATGATTTTCACTTGGAAAGTGTAACTTTGTACCTTGCAGATATGCGAACAATCCA 540
QY      548 ACCTGCTGGGCTATCTGCAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAAC 607
DB      541 ACCTGCTGGGCTATCTGCAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAAC 600
QY      608 AAGCTTACAAAACCAACCACTTCAAGACCAACCAACCAACCAACCAACCAACCAACCACT 667
DB      601 AAGCTTACAAAACCAACCACTTCAAGACCAACCAACCAACCAACCAACCAACCAACCACT 660
QY      668 AAACCAAGGAAGTATCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 727
DB      661 AAACCAAGGAAGTATCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 720
QY      728 ACACATCATCACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 787
DB      721 ACACATCATCACTACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 780
QY      788 CAATGGAAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847
DB      781 CAATGGAAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840
QY      848 ACAACATCCGAGACCCATCAACACCTCTATTTCCACCCCAACCAACCAACCAACCAAC 901
DB      841 ACAACATCCGAGACCCATCAACACCTCTATCTCCACCCCAACCAACCAACCAACCAAC 894
```

RESULT 12

US-08-801-898A-23

; Sequence 23, Application US/08801898A

; Patent No. 5998602

; GENERAL INFORMATION:

; APPLICANT: Torrence, Paul P.

; APPLICANT: Silverman, Robert H.

; APPLICANT: Cirino, Nick M.

; APPLICANT: Li, Guiying

; APPLICANT: Xiao, Wei

; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE

; TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/801,898A

; FILING DATE: 18-FEB-1997

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Poissant, Brian M.

; REGISTRATION NUMBER: 28,462

; REFERENCE/DOCKET NUMBER: 8656-009

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15222 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; FEATURE:

; NAME/KEY: RSV-A2

; LOCATION: 1...15222

; OTHER INFORMATION:

US-08-801-898A-23

Query Match 92.8%; Score 853.4; DB 2; Length 15222;

Best Local Similarity 95.5%; Pred. No. 5.4e-224;

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY      1 TGCACACATGTCACAAAACAAAGGACCAACGCGCTAAGACACTAGAAAAGACCTGGGA 60
DB      4681 TGCACACATGTCACAAAACAAAGGACCAACGCGCTAAGACACTAGAAAAGACCTGGGA 4740
QY      61 CACTCTCAATCAATTTATTTATTCATATCATCGGCTTATATAAGTTAAATCTTAAATCTGT 120
DB      4741 CACTCTCAATCAATTTATTTATTCATATCATCGTCTTATATAAGTTAAATCTTAAATCTGT 4800
QY      121 AGCACAATTCATATTCATCTTCTGGCAATGATAATCTCAACTTCACTTATTAATTAACAGC 180
DB      4801 AGCACAATTCATATTCATCTTCTGGCAATGATAATCTCAACTTCACTTATTAATTAACAGC 4860
QY      181 CATCATATTCATAGCTTCGCAACCAACCAAGTCACTAACTGAACCTGCAATCATACAAGA 240
DB      4861 CATCATATTCATAGCTTCGCAACCAACCAAGTCACTAACTGAACCTGCAATCATACAAGA 4920
QY      241 TGCACAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB      4921 TGCACAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 4980
```



Patent No. 5993824  
GENERAL INFORMATION:  
APPLICANT: Murphy, Brian R.  
APPLICANT: Collins, Peter L.  
APPLICANT: Whitehead, Stephen S.  
APPLICANT: Bukreyev, Alexander A.  
APPLICANT: Juhaz, Katalin  
TITLE OF INVENTION: PRODUCTION OF ATTENUATED RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,403A  
FILING DATE: 15-JUL-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,634  
FILING DATE: 23-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,141  
FILING DATE: 09-MAY-1997  
APPLICATION DATA:  
APPLICATION NUMBER: 17634-000510  
FILING DATE: 15-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 17634-000510  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15223 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-892-403A-1

Query Match 92.8%; Score 853.4; DB 2; Length 15223;  
Best Local Similarity 95.5%; Pred. No. 5.4e-224;  
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

1 TGCACATCTCCAAAGACGAGCCACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
4682 TGCACATCTCCAAAGACGAGCCACGCGCTTAAGACACTAGAAAGACCTGGGA 4741

61 CACTCTCAATCTATTATTATCTATCATCTCGGCTTATATAAGTTAAATCTTAAATCTGT 120  
4742 CACTCTCAATCTATTATTATCTATCATCTCGGCTTATATAAGTTAAATCTTAAATCTGT 4801

121 AGCAAAATCACAATATCTTCTGGAATGATATCTCAACTTCACTTATAATTAATACGC 180  
4802 AGCAAAATCACAATATCTTCTGGAATGATATCTCAACTTCACTTATAATTAATGAGC 4861

181 CATCATATCTAGCTCGGCAACCAACCAAGTACACTCACTCACTCATACAGAGA 240  
4862 CATCATATCTAGCTCGGCAACCAACCAAGTACACTCACTCACTCATACAGAGA 4921

241 TGCACAGCGCATCAAGACCAACCCCACTACTCTCAGGATCTCAGCTTGG 300

Db 4922 TGCACAGCGCATCAAGACCAACCCCACTACTCACCAGAAATCTCAGCTTGG 4981  
Qy 301 AATCAGCTTCTCCAAATCTGTCTGAATTTACATCAACACCAACCACTACTAGCTTCAAC 360  
Db 4982 AATCAGTCCCTCTAATCGCTGTAATTTACATCAAAATACCAATACCACTACTAGCTTCAAC 5041  
Qy 361 AACACCAGAGTCAAGTCAAACTGCAACCCACCAACAGTCAAGACTTAAATCAACCAACAA 420  
Db 5042 AACACCAGAGTCAAGTCAAACTGCAATCCAAACAGTCAAGCAAAACCAACCAACAA 5101  
Qy 421 AACCCAAACACACCCAGCAAGCCCACTACAAAACCAACCAACCAACCAACCAACCA 480  
Db 5102 AACTCAACACACACCCAGCAAGCCCACTACAAAACCAACCAACCAACCAACCAACCA 5161  
Qy 481 ACCCAATATGATTTTCACTTGAAGTCTTAACTTTGTACCTGCGCATATGAGCA 540  
Db 5162 ACCCAATATGATTTTCACTTGAAGTCTTCAACTTTGTACCTGCGCATATGAGCA 5221  
Qy 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 600  
Db 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 5281  
Qy 601 CACCAACCAAGCTTACAAAACCAACCTTCAAGCAACCAAAAGAAATCTCAAACTCA 660  
Db 5282 CACTTACCAAGCCCAACAAAACCAACCTTCAAGCAACCAAAAGAAATCTCAAACTCA 5341  
Qy 661 AACCACTAAACCAAGGAAGTACCACCAACCAAGCCCAACCAAGCAAGCAACCAACCAAC 720  
Db 5342 AACCACTAAACCAAGGAAGTACCACCAACCAAGCCCAACCAAGCAAGCAACCAACCAAC 5401  
Qy 721 CACCAAAACCAACATCACTACTGCTACCAACCAACCAACCAAGCAAGCAACCAACCAAC 780  
Db 5402 CACCAAAACCAACATCACTACTGCTACCAACCAACCAACCAAGCAAGCAACCAACCAAC 5461  
Qy 781 CACAAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTAAGCCCTTCTCA 840  
Db 5462 CACAAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTAAGCCCTTCTCA 5521  
Qy 841 AGTCTCCACACATCGAGCAGCCATCAACCTTCACTCTCCACCAACCAACCAAGCAAGCA 900  
Db 5522 AGTCTCTACACATCGAGTACCCATCAACCTTCACTCTCCACCAACCAACCAAGCAAGCA 5581  
Qy 901 GTAGTTATTAAAAA 919  
Db 5582 GTAGTTATTAAAAA 5600

RESULT 15  
US-08-720-132-1  
Sequence 1, Application US/08720132  
Patent No. 6264957  
GENERAL INFORMATION:  
APPLICANT: Collins, Peter L.  
TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY  
TITLE OF INVENTION: SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,132  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/007,083  
; FILING DATE: 27-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 15280-250-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-467-9600  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15223 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other  
US-08-720-132-1

Query Match 92.8%; Score 853.4; DB 3; Length 15223;

Best Local Similarity 95.5%; Pred. No. 5.4e-224;

Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy	1	TGCAAAACATGTCCTCAAAACAGGACCAACGCGCTAAGGACACTAGAAAAGACCTGGGA	60
Db	4682	TGCAAAACATGTCCTCAAAACAGGACCAACGCGCTAAGGACACTAGAAAAGACCTGGGA	4741
Qy	61	CACCTCTCAATCATTTATTTATTCATATCATCGGGCTATATAAGTTAAATCTTAAATCTGT	120
Db	4742	CACCTCTCAATCATTTATTTATTCATATCATCGGGCTATATAAGTTAAATCTTAAATCTGT	4801
Qy	121	AGCAAAATCACAATATTCATTCCTGGCAATGATATCTCAACTTCCTTATTAATTACAGC	180
Db	4802	AGCAAAATCACAATATTCATTCCTGGCAATGATATCTCAACTTCCTTATTAATTACAGC	4861
Qy	181	CATCATATTCATAGCTCGGCAACCAACCAAGTCACTAACAACCTGCAATCATACAAGA	240
Db	4862	CATCATATTCATAGCTCGGCAACCAACCAAGTCACTAACAACCTGCAATCATACAAGA	4921
Qy	241	TGCAAAAGCCAGATCAAGAACACAAACCCCAACATCTCACTCAGGATCTCAGCTTGG	300
Db	4922	TGCAAAAGCCAGATCAAGAACACAAACCCCAACATCTCACTCAGGATCTCAGCTTGG	4981
Qy	301	AATCAGCTTCTCCAACTGTCTGAATTTACATCAAAACCAACCACTACTAGCTTCAAC	360
Db	4982	AATCAGCTTCTCCAACTGTCTGAATTTACATCAAAACCAACCACTACTAGCTTCAAC	5041
Qy	361	AACACGAGGTCAAGTCAAACTGCAACCAACCAACCAAGTCAAGCTAATAAACAACAAC	420
Db	5042	AACACGAGGTCAAGTCAAACTGCAACCAACCAACCAAGTCAAGCTAATAAACAACAAC	5101
Qy	421	AACCCAAACACAAACCCAGCAAGCCCTACTACAAACCAAGCCCAAAACCAACCAACAA	480
Db	5102	AACCCAAACACAAACCCAGCAAGCCCTACTACAAACCAAGCCCAAAACCAACCAACAA	5161
Qy	481	ACCCAAATATGATTTTCATCTCGAAGTGTAACTTTGTACCTGCGAGCATATGCGACAA	540
Db	5162	ACCCAAATATGATTTTCATCTCGAAGTGTAACTTTGTACCTGCGAGCATATGCGACAA	5221
Qy	541	CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACAAACCAAGGAAAGAAAC	600
Db	5222	CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACAAACCAAGGAAAGAAAC	5281
Qy	601	CACCACCAAGCCTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA	660
Db	5282	CATTACCAAGCCCAAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA	5341
Qy	661	AACCACTAAACCAAGGAGTACCCACCAAGCCCAAGAGGACCAACCATCAACAC	720
Db	5342	AACCACTAAACCAAGGAGTACCCACCAAGCCCAAGAGGACCAACCATCAACAC	5401
Qy	721	CACCAAAACAAACATCACTACTGCTCAACCAACCAACCAAGGAAATCCAAACT	780
Db	5402	CACCAAAACAAACATCACTACTGCTCAACCAACCAACCAAGGAAATCCAAACT	5461

Qy	781	CACAAGTCAAAATGGAACCTTCCACTCAACTCTCCAGGCAATCTTAAGCCCTTCTCA	840
Db	5462	CACAAGTCAAAATGGAACCTTCCACTCAACTCTCCAGGCAATCTTAAGCCCTTCTCA	5521
Qy	841	AGTCTCCACAACATCCGAGCACCCATCACAACCTCTATCTCCACCAACCAACACGCA	900
Db	5522	AGTCTCTACAACATCCGAGTACCCATCACAACCTTCTATCTCCACCAACCAACACGCA	5581
Qy	901	GTAAGTTATTAATAAAAAA 919	
Db	5582	GTAAGTTATTAATAAACATA 5600	

Search completed: October 30, 2003, 01:17:36  
Job time : 74.4618 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:50:52 ; Search time 313.982 Seconds  
(without alignments)  
7909.644 Million cell updates/sec

Title: US-09-462-816-1  
Perfect score: 920  
Sequence: 1 tgcacacatgtccaaaaca.....gtagttattataaaaaaaa 920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
	2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
	3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
	4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
	5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
	6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
	7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
	8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
	9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
	10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
	11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
	12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
	13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
	14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
	15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
	16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
	17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
	18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
	19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
	20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
	21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
	22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
	23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
	24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
	25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920	100.0	920	14	AAQ45686 Respiratory syncyt
2	920	100.0	920	20	AAQ45686 Membrane bound G p
3	857.6	93.2	935	8	AAV18736 Sequence encoding
4	857.6	93.2	935	13	AAV18736 HRSV glycoprotein
5	856	93.0	935	15	AAQ29623 HRSV glycoprotein
6	853.4	92.8	15222	18	AAV18440 Human respiratory
7	853.4	92.8	15223	18	AAV18440 Respiratory syncyt
8	853.4	92.8	15223	19	AAV17553 Respiratory syncyt

9	853.4	92.8	15223	21	AAV88743 Respiratory syncyt
10	848.2	92.2	918	22	AAV88743 Human RSV G-prote
11	837.8	91.1	897	24	AAV88743 Respiratory syncyt
12	830.4	90.3	15210	20	AAV88743 Polynucleotide seq
13	715	77.7	715	20	AAV88743 G protein gene fra
14	637.4	69.3	696	19	AAV88743 Respiratory syncyt
15	441.4	48.0	15225	19	AAV88743 Respiratory syncyt
16	441.4	48.0	15225	21	AAV88743 Respiratory syncyt
17	441.4	48.0	15225	24	AAV88743 Human respiratory
18	422.2	45.9	15229	19	AAV88743 RSV isolate 18537
19	422.2	45.9	15229	20	AAV88743 Nucleotide sequenc
20	422.2	45.9	15229	20	AAV88743 DNA encoding the L
21	411.4	44.7	15218	19	AAV88743 RSV isolate 2B wil
22	411.4	44.7	15218	20	AAV88743 Nucleotide sequenc
23	411.4	44.7	15218	20	AAV88743 Nucleotide sequenc
24	411.4	44.7	15218	20	AAV88743 DNA encoding the L
25	411.4	44.7	15219	19	AAV88743 RSV vaccine 2B33F
26	411.4	44.7	15219	19	AAV88743 RSV vaccine 2B20L
27	411.4	44.7	15219	19	AAV88743 RSV revertant 2B33
28	411.4	44.7	15219	19	AAV88743 RSV revertant 2B20
29	411.4	44.7	15219	20	AAV88743 Nucleotide sequenc
30	411.4	44.7	15219	20	AAV88743 Nucleotide sequenc
31	411.4	44.7	15219	20	AAV88743 Nucleotide sequenc
32	411.4	44.7	15219	20	AAV88743 DNA encoding the L
33	411.4	44.7	15219	20	AAV88743 DNA encoding the L
34	411.4	44.7	15219	20	AAV88743 DNA encoding the L
35	411.4	44.7	15219	20	AAV88743 DNA encoding the L
36	212.8	23.1	1050	17	AAV88743 Encodes Streptococ
37	212.8	23.1	1050	22	AAV88743 Nucleotide sequenc
38	212.8	23.1	1356	22	AAV88743 Nucleotide sequenc
39	210.8	22.9	303	16	AAV88743 RSV subgroup A clon
40	210.8	22.9	303	17	AAV88743 RSV subgp. A prote
41	210.8	22.9	303	17	AAV88743 RSV sub-group A wi
42	210.8	22.9	303	17	AAV88743 Respiratory Syncyt
43	210.8	22.9	303	20	AAV88743 RSV G protein anti
44	210.8	22.9	303	21	AAV88743 DNA encoding a G2N
45	210.8	22.9	303	22	AAV88743 Nucleotide sequenc

ALIGNMENTS

RESULT 1  
AAQ45686  
ID AAQ45686 standard; DNA; 920 BP.

XX	AAQ45686;
XX	25-MAR-2003 (updated)
DT	13-JAN-1994 (first entry)
XX	Respiratory syncytial virus (RSV) G gene.
DE	PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine;
KW	ds.
XX	Respiratory syncytial virus.
OS	XX
XX	XX
PH	Key Location/Qualifiers
FT	CDS 8...901
FT	/*tag= a
FT	/product= RSV G protein
FT	89..176
FT	/*tag= b
FT	/label= Transmembrane anchor domain
XX	XX
XX	WO9314207-A1.
PN	22-JUL-1993.
PD	XX
XX	XX
XX	05-JAN-1993; 93WO-CA00001.
XX	06-JAN-1992; 92GB-0000117.
PR	XX



XX (CONN-) CONNAUGHT LAB LTD.  
 XX PI Ewasysbyn ME, Klein MH;  
 XX DR WPI; 1993-243222/30.  
 XX DR P-PSDB; AAR39286.  
 XX  
 PT Multimeric hybrid genes and their chimeric proteins - are  
 PT vaccines against multiple pathogenic infections e.g.  
 PT para-influenza virus and respiratory syncytial virus  
 XX  
 PS Claim 11; Figure 7A-7D; 80pp; English.  
 XX  
 CC A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;  
 Query Match 100.0%; Score 920; DB 14; Length 920;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e-211;  
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACACATGCTCCAAACCAAGCAGCAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 Db 1 TGCACACATGCTCCAAACCAAGCAGCAGCAGCGCTAGACACTAGAAAGACCTGGGA 60  
 QY 61 CACTCTCAATCATTTATTATTATCATATCATCGGCTTATATAGTTAAATCTTAAATCTGT 120  
 Db 61 CACTCTCAATCATTTATTATTATCATATCATCGGCTTATATAGTTAAATCTTAAATCTGT 120  
 QY 121 AGCACAATACATATTCATTTCTGGCAATGATAATCTCAACTTCACTTATATATTACAGC 180  
 Db 121 AGCACAATACATATTCATTTCTGGCAATGATAATCTCAACTTCACTTATATATTACAGC 180  
 QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTTAAGACTGCAATCATACAGA 240  
 Db 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTTAAGACTGCAATCATACAGA 240  
 QY 241 TGCACAAGCCAGATCAAGAACCAACACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 Db 241 TGCACAAGCCAGATCAAGAACCAACACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 QY 301 AATCAGCTTCTCAATCTGTCTGAATTTATCATCAAAACCAACCCATAGCTTCAAC 360  
 Db 301 AATCAGCTTCTCAATCTGTCTGAATTTATCATCAAAACCAACCCATAGCTTCAAC 360  
 QY 361 AACACGAGGATCAAGTCAAACTCGCAACCAACCAAGTCACAGTCAAGACTTAAACACACAC 420  
 Db 361 AACACGAGGATCAAGTCAAACTCGCAACCAACCAAGTCACAGTCAAGACTTAAACACACAC 420  
 QY 421 AACCCAAACACACCCAGCAGCCACTACAAACCAACGCGCAAAACCAACCAACCAACAA 480  
 Db 421 AACCCAAACACACCCAGCAGCCACTACAAACCAACGCGCAAAACCAACCAACCAACAA 480  
 QY 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTACCTCGCAGCATATGCGCAA 540  
 Db 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTACCTCGCAGCATATGCGCAA 540  
 QY 541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAAAACCAAGCAAGAAAC 600  
 Db 541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAAAACCAAGCAAGAAAC 600  
 QY 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660  
 Db 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660

QY 661 AACCACTAAACCAAGGAAGTACCCACCAACCAAGGCCACAGAAAGCCCAACCATCAACAC 720  
 Db 661 AACCACTAAACCAAGGAAGTACCCACCAACCAAGGCCACAGAAAGCCCAACCATCAACAC 720  
 QY 721 CACCAAAACCAACATCAACACTACACTGCTCACCACCAACCAACAGGAAATCCAAACT 780  
 Db 721 CACCAAAACCAACATCAACACTACACTGCTCACCACCAACCAACAGGAAATCCAAACT 780  
 QY 781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840  
 Db 781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840  
 QY 841 AGTCTCCCAACATCCGAGCAGCCCATCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 900  
 Db 841 AGTCTCCCAACATCCGAGCAGCCCATCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 900  
 QY 901 GTAGTTATTAAAAA 920  
 Db 901 GTAGTTATTAAAAA 920

RESULT 2  
 AAX08421  
 ID AAX08421 standard; cDNA; 920 BP.  
 XX  
 AC AAX08421;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein gene of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator; ss.  
 XX  
 OS Respiratory syncytial virus (RSV).  
 XX

Key Location/Qualifiers  
 CDS 8..904  
 FT /\*tag= a  
 FT /product= "Membrane bound G protein"  
 XX  
 PN WO9904010-A1.  
 PD 28-JAN-1999.  
 XX  
 PF 16-JUL-1998; 98MO-CA00697.  
 XX  
 PR 18-JUL-1997; 97US-0896442.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Klein MH, Li X, Sambhara S;  
 XX  
 DR WPI; 1999-132254/11.  
 DR P-PSDB; AAW96313.  
 XX

Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 3; Figure 2; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when

CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.

XX SQ Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;  
Query Match 100.0%; Score 920; DB 20; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.4e-211;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGC AAAAATGTCGCAAAACAGAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60  
Db 1 TGC AAAAATGTCGCAAAACAGAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60

Qy 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
Db 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120

Qy 121 AGCAAAATCACAATATCCATTTCTGGCAATGATATCTCAACTTCACTTAAATTAACAGC 180  
Db 121 AGCAAAATCACAATATCCATTTCTGGCAATGATATCTCAACTTCACTTAAATTAACAGC 180

Qy 181 CATCATATTCATAGCTCGGCAAAACCAAGTCACTAACAACCTGCAATCATACAAGA 240  
Db 181 CATCATATTCATAGCTCGGCAAAACCAAGTCACTAACAACCTGCAATCATACAAGA 240

Qy 241 TGC AACAGCAGATCAAGAACCAACCCCAACATCTCACTCAGGATCTCAGCTTGG 300  
Db 241 TGC AACAGCAGATCAAGAACCAACCCCAACATCTCACTCAGGATCTCAGCTTGG 300

Qy 301 AATCAGCTTCTCCAAATCTGTCTGAATTTATACATCAAAACCAACCACTACTAGCTTCAAC 360  
Db 301 AATCAGCTTCTCCAAATCTGTCTGAATTTATACATCAAAACCAACCACTACTAGCTTCAAC 360

Qy 361 AACACAGAGGTCAAGTCAAACTGCAACCCCAACATCTCACTCAGGATCTCAGCTTGG 420  
Db 361 AACACAGAGGTCAAGTCAAACTGCAACCCCAACATCTCACTCAGGATCTCAGCTTGG 420

Qy 421 AACCCAAACCAACCCCAAGCCCACTACAAACAGCCCAACCAACCAACCAACAA 480  
Db 421 AACCCAAACCAACCCCAAGCCCACTACAAACAGCCCAACCAACCAACCAACAA 480

Qy 481 ACCCAATATGATTTTCACTTTCGAAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
Db 481 ACCCAATATGATTTTCACTTTCGAAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540

Qy 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGAAAGAAAC 600  
Db 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGAAAGAAAC 600

Qy 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660  
Db 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACCAACCAAAAGATCTCAAACTCA 660

Qy 661 AACCACTAAACCAAGGAAGTACCCACCAAGAGCCCAAGAGAGGACCAACCAACAC 720  
Db 661 AACCACTAAACCAAGGAAGTACCCACCAAGAGCCCAAGAGAGGACCAACCAACAC 720

Qy 721 CACCAAAAACCAATCAACTACATGCTGCAACCAACCAACCAAGGAATCCAAACT 780  
Db 721 CACCAAAAACCAATCAACTACATGCTGCAACCAACCAACCAAGGAATCCAAACT 780

Qy 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840

Qy 841 AGTCTCCACCAATCCGAGCAGCCATCACAACCTCTATCTCCACCCCAACACAGCCCA 900  
Db 841 AGTCTCCACCAATCCGAGCAGCCATCACAACCTCTATCTCCACCCCAACACAGCCCA 900

Qy 901 GTAGTTATTAAAAA 920  
|||||

Db 901 GTAGTTATTAAAAA 920  
AAAT70784  
ID AAN70784 standard; cDNA; 935 BP.  
XX AC AAN70784;  
XX AC AAN70784;  
DT 25-MAR-2003 (updated)  
XX 05-APR-1991 (first entry)  
DE Sequence encoding human respiratory syncytial virus (HRSV) A2 strain  
DE G protein.  
XX Vaccine; ss.  
XX Human respiratory syncytial virus (HRSV).  
XX Key Location/Qualifiers  
XX CDS 16..913  
XX /\*tag= a  
XX MO8704185-A.  
XX PN PN  
XX 16-JUL-1987.  
XX 23-DEC-1986; 86WO-US02756.  
XX 14-JAN-1986; 86US-0818740.  
XX (UYN-) UNIV NORTH CAROLINA.  
XX (WERTZ) WERTZ G W.  
DR WPI; 1987-206300/29.  
DR P-PSDB; AAP70845.  
XX Vaccines for human respiratory virus - comprising proteins or  
XX fragment encoded by a DNA sequence coding for human respiratory  
XX syncytial virus proteins.  
XX Disclosure; Chart 13; 57pp; English.  
XX A novel plasmid which comprises a DNA sequence encoding this  
XX protein, and the protein itself, are claimed, for use as HRSV  
XX vaccines. The vaccine can be administered to pregnant women or to  
XX women of child bearing age to stimulate maternal antibodies.  
XX Infants can also be vaccinated at 2-3 months of age.  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;  
Query Match 93.2%; Score 857.6; DB 8; Length 935;  
Best Local Similarity 95.8%; Pred. No. 1.4e-196;  
Matches 881; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 TGC AAAACATGTCTCAAAAACAGAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60  
Db 9 TGC AAAACATGTCTCAAAAACAGAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 68

Qy 61 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
Db 69 CACTCTCAATCATTTATTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 128

Qy 121 AGCAAAATCACAATATCCATTTCTGGCAATGATATCTCAACTTCACTTAAATTAACAGC 180  
Db 129 AGCAAAATCACAATATCCATTTCTGGCAATGATATCTCAACTTCACTTAAATTAACAGC 188

Qy 181 CATCATATTCATAGCTCGGCAAAACCAAGTCACTTAAATTAACAGC 240  
Db 189 CATCATATTCATAGCTCGGCAAAACCAAGTCACTTAAATTAACAGC 248

Qy 241 TGC AACAGCAGATCAAGAACCAACCCCAACATCTCACTCAGGATCTCAGCTTGG 300



```

Db 609 CACTACCAAGCCCAAAAAACCAACCCCTCAAGACCAACCAAAAAAGATCCCAACCTCA 668
QY 661 AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCATCAACAC 720
Db 669 AACCACTAAATCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCATCAACAC 728
QY 721 CACCAAAACCAACATCAACTACACTGCTCACCACCAACCAACCAAGGAATCCAAACT 780
Db 729 CACCAAAACCAACATCAACTACACTGCTCACCACCAACCAACCAAGGAATCCGAACT 788
QY 781 CACAGTCAATGAACCTTCCTCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
Db 789 CACAAGTCAATGAACCTTCCTCACTCAACCTCTCCGAAGGCAATCCAGCCCTTCTCA 848
QY 841 AGTCTCCACACATCCCGAGGACCCATCAACACCTCTATCCACCCCAACCAAGCCCA 900
Db 849 AGTCTCTACACATCCCGAGGATCCCATCAACACCTCTATCCACCCCAACCAAGCCCA 908
QY 901 GTAGTTATTAAAAA 920
Db 909 GTAGTTATTAAAAA 928

RESULT 5
AAQ29623
ID AAQ29623 standard; DNA; 935 BP.
XX
AC AAQ29623;
XX
DT 25-MAR-2003 (updated)
DT 03-MAR-1993 (first entry)
XX
DE HSRV glycoprotein G (gpG).
XX
KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KW major capsid protein; N; ss.
XX
OS Human respiratory syncytial virus strain A2.
XX
FH Key
FT CDS 16..912
FT Location/Qualifiers
FT /*tag= a
FT /label= G_protein
FT misc_feature 16..22
FT /*tag= b
FT /note= "oligonucleotide used to probe for
FT full length cDNA"
FT misc_feature 898..912
FT /*tag= c
FT /note= "oligonucleotide used to specifically
FT prime the reverse transcription reaction
FT for making the first strand of the cDNA"
XX
PN US5149650-A.
XX
PD 22-SEP-1992.
XX
PF 13-JUL-1988; 88US-0218737.
XX
PR 14-JAN-1986; 86US-0818740.
XX 13-JUL-1988; 88US-0218737.
XX (UYN-) UNIV NORTH CAROLINA.
XX
PI Collins PL, Wertz GW;
XX
DR WPI; 1992-340247/41.
DR P-PSDB; AAR25302.
XX
PT Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments

```

```

XX Disclosure; Page 18; 2lpp; English.
XX
CC The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;

Query Match 93.0%; Score 856; DB 13; Length 935;
Best Local Similarity 95.7%; Pred. No. 3.4e-196;
Matches 880; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1 TGCACCAATGTCACCAAAACCAAGGACCAACGACCGCTAGACACTAGAAAGACCTGGGA 60
Db 9 TGCACCAATGTCACCAAAACCAAGGACCAACGACCGCTAGACACTAGAAAGACCTGGGA 68
QY 61 CACTCTCAATCATTTATTATTATCATATCATGCTGCTTATATAGTTAAATCTGT 120
Db 69 CACTCTCAATCATTTATTATTATCATATCATGCTGCTTATATAGTTAAATCTGT 128
QY 121 AGCACAATCACTATTCATTCTGCAATGATATCTCACTCTTATTAATTAACAGC 180
Db 129 AGCACAATCACTATTCATTCTGCAATGATATCTCACTCTTATTAATTAACAGC 188
QY 181 CATCATATTTCATAGCTCGGCAACCAACAAAGTCACACTAAGCAATGCAATCAAGA 240
Db 189 CATCATATTTCATAGCTCGGCAACCAACAAAGTCACACTAAGCAATGCAATCAAGA 248
QY 241 TGCACCAAGCCAGATCAAGAACCAACCCCAATACCTCACTCAGGATCCTCAGCTGG 300
Db 249 TGCACCAAGCCAGATCAAGAACCAACCCCAATACCTCAGGATCCTCAGCTGG 308
QY 301 AATCAGCTCTTCCATCTCTGTAATTAATCACTAAGCAACCAACCATAGCTTCAAC 360
Db 309 AATCAGCTCTTCCATCTCTGTAATTAATCACTAAGCAACCAACCATAGCTTCAAC 368
QY 361 AACACCAAGGATCAAGTCAAACTCGAACCCCAACAGTCAAGCTAAGAAACACACACAC 420
Db 369 AACACCAAGGATCAAGTCAAACTCGAACCCCAACAGTCAAGCTAAGAAACACACACAC 428
QY 421 AACCCCAACCAACCCCAAGCCCACTAACAACCAACCCCAACCAACCAACCAACCAAC 480
Db 429 AACTCAAAACCAACCCCAAGCCCACTAACAACCAACCCCAACCAACCAACCAACCAAC 488
QY 481 ACCCAATATGATTTTCTTCAAGTGTGTTAACTTTGTAACCTGAGCATATGACGAA 540
Db 489 ACCCAATATGATTTTCTTCAAGTGTGTTAACTTTGTAACCTGAGCATATGACGAA 548
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATCAACCAACCAACCAACCAACCAACCAAC 600
Db 549 CAATCCAACTGCTGGGCTATCTGCAAAAGATATCAACCAACCAACCAACCAACCAACCAAC 608
QY 601 CACCAACCAAGCTCAAAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 660
Db 609 CACTACCAAGCCCAACCAACCAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 668
QY 661 AACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCATCAACAC 720
Db 669 AACCACTAAATCAAGGAAGTACCCACCAAGCCCAAGAGCCCAACCATCAACAC 728
QY 721 CACCAAAACCAACATCAACTACACTGCTCACCACCAACCAACCAAGGAATCCGAACT 780
Db 729 CACCAAAACCAACATCAACTACACTGCTCACCACCAACCAACCAAGGAATCCGAACT 788
QY 781 CACAAGTCAATGAAGAACTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
Db 789 CACAAGTCAATGAAGAACTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 848
QY 841 AGTCTCTCAACATCCGAGGACCCATCAACACCTCTATCCACCCCAACCAAGCCCA 900

```

Db	849	AGTCTCTCAACATCCGAGTACCCATCACACCTTCTCATCTCCACCCACACACACGCGCA	908
Qy	901	GTAGTATTATTAACAAAAA 920	
Db	909	GTAGTTACTTAAAAA 928	
RESULT 6			
AAT78440			
ID	AAT78440 standard; DNA; 15222 BP.		
XX	AC	AAT78440;	
XX	17-MAR-1998	(first entry)	
XX	Human respiratory syncytial virus strain A2.		
XX	Antisense oligonucleotide; respiratory syncytial virus;		
KW	RSV; treatment; infection; inhibition; strain A2; ss.		
XX	Human respiratory syncytial virus.		
OS	WO9729757-A1.		
XX	21-AUG-1997.		
XX	14-FEB-1997;	97WO-US02531.	
XX	15-FEB-1996;	96US-0011725.	
PR	(CLEV-) CLEVELAND CLINIC FOUND.		
PA	(USSH ) US NAT INST OF HEALTH.		
XX	Cirino NM, Li G, Silverman RH, Torrence PF, Xiao W;		
PI	WPI; 1997-424748/39.		
XX	Polynucleotide containing sequence anti-sense to region of RSV -		
XX	connected via a linker to an activator of RNaseL, used to treat RSV		
PT	infections		
XX	Disclosure; Pages 47-51; 89pp; English.		
PS	The present sequence was used in the preparation of a novel		
CC	polynucleotide, comprising an antisense oligonucleotide, with a		
CC	hydroxy group at one end, that is complementary to 15-20 bases of		
CC	the anti-genomic RNA strand of a respiratory syncytial virus (RSV),		
CC	a linker attached to the OH-end of the antisense oligonucleotide and		
CC	an oligonucleotide activator of RNaseL attached to the linker. The		
CC	polynucleotide can be used to treat RSV infections, which can also		
CC	be treated by administration of the antisense oligonucleotide, so as		
CC	to form a complex with activated RNase L in vivo. The		
CC	polynucleotide can be transported across the cell membranes without		
CC	carriers or permeability agents, and once introduced destroys		
CC	antisense target RNA. It also inhibits RSV infection in vitro in a		
CC	superior manner to the conventional drug, ribavirin.		
XX	Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;		
SQ			
Query Match 92.8%; Score 853.4; DB 18; Length 15222;			
Best Local Similarity 95.5%; Pred. No. 3.2e-195;			
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0			
Qy	1	TGCAACATGTCCAAAACAGGACCAACGACCGCTTAAGCATTAGAAAGACCTGGGA	60
Db	4681	TGCAACATGTCCAAAACAGGACCAACGACCGCTTAAGCATTAGAAAGACCTGGGA	474
Qy	61	CACCTCTCAATCATTTTATTATTCATATCATCGGGCTTATATAAGTTAAATCTGT	120
Db	4741	CACCTCTCAATCATTTTATTATTCATATCATCGGGCTTATATAAGTTAAATCTGT	480
Qy	121	AGCAACAATCACATTTATTCATCTCGGCAATGATTAATCTCACTTCACTTATAATTACAGC	180



CC	Gln831, Met169 or Tyr1321 in the RSV polymerase gene or a ts
CC	nucleotide substitution in the gene-start sequence of gene M2.
CC	Also claimed are: (1) an isolated infectious RSV particle which
CC	comprises a recombinant RSV (anti)genome, N, P, and L proteins, a
CC	RNA polymerase elongation factor, where the (anti)genome is modified:
CC	(1) to ablate or modulate expression of a SH, NS1, NS2 or G gene or
CC	a cis-acting regulatory sequence; and (ii) by a termination codon
CC	introduced within a selected gene, or by a change in sequence,
CC	position or presence of a GS or GE transcription signal relative to
CC	the selected gene; (2) an expression vector; and (3) an RSV strain
CC	selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR
CC	2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2455), or
CC	cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or
CC	B-1 cp52/B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated
CC	attenuated recombinant RSV and RSV particles are used in a vaccine
CC	to stimulate the immune system of an individual to induce
CC	protection against RSV. The expression vector of (2) is used for
CC	the production of infectious attenuated RSV particles.
XX	
SQ	Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;
	Query Match 92.8%; Score 853.4; DB 19; Length 15223;
	Best Local Similarity 95.5%; Pred. No. 3.2e-195;
	Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY	1 TGCAACATGTCGAATAAACAAGGACCAACGCACCGCTAAGACACTAGAAGAAGACCTGGGA 60
DB	4682 TGCAACATGTCGAATAAACAAGGACCAACGCACCGCTAAGACACTAGAAGAAGACCTGGGA 4741
QY	61 CACTCTCAATCATTTATTATTTCATATCATCGCGCTTATATAAGTTAAATCTTAAATCTGT 120
DB	4742 CACTCTCAATCATTTATTATTTCATATCATCGTGCTTATATAAGTTAAATCTTAAATCTGT 4801
QY	121 AGCACAATCACATPATTCATTCCTGGCAATGATAATCTCAACTTCATTAATTAACAGC 180
DB	4802 AGCACAATCACATPATTCATTCCTGGCAATGATAATCTCAACTTCATTAATTAATTCAGC 4861
QY	181 CATCATATTTATAGCTCTGGGCAACCACRAAGTCACACTAACACTGCAATCATACAAGA 240
DB	4862 CATCATATTTATAGCTCTGGGCAACCACRAAGTCACACTAACACTGCAATCATACAAGA 4921
QY	241 TGCAACAAGCCAGATCAAGAACAACGCCCAAATACTCTCACTCAGGATCCTCAGCTTGG 300
DB	4922 TGCAACAAGCCAGATCAAGAACAACGCCCAAATACTCTCAGGATCCTCAGCTTGG 4981
QY	301 AATCAGTCTTCCAAATCTGTCTGAAATTAATATCAAAAACCAACCATATCTAGCTTCAAC 360
DB	4982 AATCAGTCTTCCAAATCTGTGAAATTAATATCAAAAACCAACCATATCTAGCTTCAAC 5041
QY	361 AACCAACGAGGTCAAGTCMAACCTGCAACCCCAACAACAGTCAAGACTAAAAACAACAAC 420
DB	5042 AACCAACGAGGTCAAGTCAACCTGCAATCCCAACACAGTCAAGACCAAAAACAACAAC 5101
QY	421 AACCCAACAACAACCCAGACGCCACTTACAAAAACAACGCCCAAAACCAACCAACAAC 480
DB	5102 AACTCAACAACAACCCAGACGCCCAACCAACAACAACGCCCAAAACCAACCAACAAC 5161
QY	481 ACCCAATATGATTTGACTTCGAGAGTGTTTAACTTTGTACCTTGCAGCATATGCAGCAA 540
DB	5162 ACCCAATATGATTTTCACTTTGAGTGTTTCAACTTTGTACCTTGCAGCATATGCAGCAA 5221
QY	541 CAATCCACCTGCTGGGCTATCTGCAAGAATACCAAAACCAAAAACCCAGGAAGAAAC 600
DB	5222 CAATCCACCTGCTGGGCTATCTGCAAGAATATCCAAAACCAAAAACCCAGGAAGAAAC 5281
QY	601 CACCACCAAGCTTACAAAAAACCAACCTTTCAGACCAACCAAAAAAGATCTTCAAACTTCA 660
DB	5282 CACTACCAAGCCCCAABAAAACCAACCTTTCAGACCAACCAAAAAAGATCCCAAACTTCA 5341
QY	661 AACCACTAAACAAGGAAGTAGTCCCACACCAAGCCGACAGAGAGCCACCATCAACAC 720
DB	5342 AACCACTAAATCAAGGAAGTAGTCCCACCAACCAAGCCCAAGAGAGCCACCATCAACAC 5401

Qy	721	CACCAAAACAAACATCACAACTACACTGCTCCACCAACACACACAGGAATCCAAAAC	780
Db	5402	CACAAACAAACATCATACTACTACTACTCTCACTCCAAACCCACAGGAATCCAGACT	5461
Qy	781	CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA	840
Db	5462	CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCCAAAGCCCTTCTCA	5521
Qy	841	AGTCTCACAACATCCGAGCAGCCATCAGAACCTCTATCTCCACCCCAACACACAGCCCA	900
Db	5522	AGTCTCTACAAATCCGAGTACCCATCACAACCTTCACTCTCCACCCCAACACACAGCCCA	5581
Qy	901	GTAGTTATTAATAAAAAA 919	
Db	5582	GTAGTTATTAATAAACATA 5600	
RESULT 9			
AAA88743			
ID	AAA88743	standard; cDNA; 15223 BP.	
XX	XX	AAA88743;	
XX	XX	19-FEB-2001 (first entry)	
XX	XX	Respiratory syncytial virus D46 5'-3' positive sense sequence.	
XX	XX	RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.	
XX	XX	Chimeric - Human respiratory syncytial virus.	
OS	OS	Chimeric - Bacteriophage T7.	
OS	OS	Synthetic.	
XX	XX	Key	Location/Qualifiers
FT	FT	variation	replace(4,G)
FT	FT	mutation	/*tag= a
FT	FT		1099
FT	FT		/*tag= b
FT	FT		/note= "single C insertion to create AflII site
FT	FT		in the NS2-N intergenic region"
FT	FT	mutation	replace(1139,A)
FT	FT		/*tag= C
FT	FT		/note= "Creates NcoI site in N gene nontranslated
FT	FT		region"
FT	FT	mutation	replace(1140,G)
FT	FT		/*tag= d
FT	FT		/note= "Creates NcoI site in N gene nontranslated
FT	FT		region"
FT	FT	mutation	replace(5612,A)
FT	FT		/*tag= e
FT	FT		/note= "creates StuI site in G/F intergenic region"
FT	FT	mutation	replace(5616,A)
FT	FT		/*tag= f
FT	FT		/note= "creates StuI site in G/F intergenic region"
FT	FT	mutation	replace(7560,A)
FT	FT		/*tag= g
FT	FT		/note= "creates SphI site in F/M2 intergenic region"
XX	XX	WO2000061611-A2.	
XX	XX	19-OCT-2000.	
XX	XX	31-MAR-2000; 2000WO-US08802.	
XX	XX	13-APR-1999; 99US-0291894.	
XX	XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
XX	XX	Collins PL, Murphy BR, Whitehead SS;	
XX	XX	WPI; 2000-679462/66.	
XX	XX	Infectious chimeric respiratory syncytial virus (RSV) produced from	



cloned nucleotide sequences, useful as a vaccine against diseases caused by the virus, such as pneumonia and bronchiolitis -  
 Example 7; Page 262-268; 280pp; English.

The present sequence is that of the 5' to 3' positive-sense sequence of human respiratory syncytial virus (RSV); the genome itself is negative-sense. This antigenome cDNA, termed D46, was synthesized in segments by RT-PCR using synthetic oligonucleotides as primers and intracellular RSV mRNA or genome RNA isolated from purified viruses as template. The antigenome includes a 5'-terminal nonviral G triplet contributed by the T7 promoter, 4 sequence markers (see AAA88745-47) at positions 1099 (which adds 1 nucleotide to the length), 1139, 5611 and 7559, a ribozyme and tandem T7 terminators, and a single nonviral 3'-phosphorylated residue contributed to the 3' end by ribozyme cleavage. The invention provides an isolated infectious chimeric RSV comprising a major nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a large polymerase protein (L), an RNA polymerase elongation factor, and a partial or complete RSV genome or antigenome of one RSV strain or subgroup virus combined with a heterologous gene of a different RSV strain or subgroup virus. The chimeric RSV is infectious and attenuated, preferably by introduction of selected mutations. It is useful as a vaccine against RSV, which causes diseases such as pneumonia and bronchiolitis in infants. The immune system of an individual is stimulated to induce protection against natural RSV infection, preferably in a multivalent manner to achieve protection against multiple RSV strains and/or subgroups.

Sequence 15223 BP; 5921 A; 2705 C; 2361 G; 4236 T; 0 other;

Query Match 92.8%; Score 853.4; DB 21; Length 15223;  
 Best Local Similarity 95.5%; Pred. No. 3.2e-195;  
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

1 TGCACATGTCCTCAAAACAAAGACACAAACGACGCTTAAGACACTAGAAAGACCTGGGA 60  
 4682 TGCACATGTCCTCAAAACAAAGACACAAACGACGCTTAAGACACTAGAAAGACCTGGGA 4741  
 61 CACTCTCAATCAATTAATTCATATCATCATCGGCTTATATTAAGTTAAATCTTAATCTGT 120  
 4742 CACTCTCAATCAATTAATTCATATCATCATCGGCTTATATTAAGTTAAATCTTAATCTGT 4801  
 121 AGCACAATCAATTCATTCATTCCTGGCAATGATTAATCTCAATTCATTAATTCATTCAGC 180  
 4802 AGCACAATCAATTCATTCATTCCTGGCAATGATTAATCTCAATTCATTAATTCATTCAGC 4861  
 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACTGCAATCATACAGA 240  
 4862 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACTGCAATCATACAGA 4921  
 241 TGCAACAAGCCAGATCAAGAACACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 300  
 4922 TGCAACAAGCCAGATCAAGAACACAAACCCCAACATACCTCACTCAGGATCCTCAGCTTGG 4981  
 301 AATCAGTCTTCAATCTGCTGCAATTCATCATCAACACACACATACCTCACTCAGCTTGG 360  
 4982 AATCAGTCTTCAATCTGCTGCAATTCATCATCAACACACATACCTCACTCAGCTTGG 5041  
 361 AACACACAGGATCAAGTCAAACTGCAACCCCAACCAAGTCACACTAACTGCAATTCATCAAC 420  
 5042 AACACACAGGATCAAGTCAAACTGCAACCCCAACCAAGTCACACTAACTGCAATTCATCAAC 5101  
 421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACCAACCGCAAAACCAACCAACCAACAA 480  
 5102 AACTCAACACAAACCCAGCAAGCCCACTACAAACCAACCGCAAAACCAACCAACCAACAA 5161  
 481 ACCCAATATGATTTTCACTTCGAGAGTGTATCTTTGATCTTACCTCGAGCATATGCGACAA 540  
 5162 ACCCAATATGATTTTCACTTCGAGAGTGTATCTTTGATCTTACCTCGAGCATATGCGACAA 5221  
 541 CAAATCAACCTGCTGGGCTATCTGCAAAAGAAATACCAACCAAAACCAAGCAAGAAAC 600

Db 5222 CAAATCAACCTGCTGGGCTATCTGCAAAAGAAATACCAACCAAAACCAAGCAAGAAAC 5281  
 QY 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTTCAAACTCA 660  
 Db 5282 CACTACCAAGCCCAACAAAAAACCAACCTTCAAGACAAACCAAAAAAGATCTTCAAACTCA 5341  
 QY 661 AACCACTAAACCAAGCAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 720  
 Db 5342 AACCACTAAATCAAGCAAGTACCCACCAAGCCCAAGAGCCCAACCAATCAACAC 5401  
 QY 721 CACCAAAACCAACATCACAACTACATCTGCTCACCACCAACCAAGGAAATCCAAACT 780  
 Db 5402 CACCAAAACCAACATCACAACTACATCTGCTCACCACCAACCAAGGAAATCCAAACT 5461  
 QY 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGCAATCTTAAGCCCTTCTCA 840  
 Db 5462 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGCAATCTTAAGCCCTTCTCA 5521  
 QY 841 AGTCTCCCAACATCCGACGACCCCATCAACCTCTCAACCTCTCAACCTCAACCAAGCCCA 900  
 Db 5522 AGTCTCTCAACATCCGACGACCCCATCAACCTCTCAACCTCTCAACCTCAACCAAGCCCA 5581  
 QY 901 GTAGTTATTAAAAA 919  
 Db 5582 GTAGTTATTAAAAACATA 5600  
 RESULT 10  
 AAC88494  
 ID AAC88494 standard; DNA; 918 BP.  
 AC AAC88494;  
 DT 12-MAR-2001 (first entry)  
 XX Human RSV G-protein gene.  
 XX Chimeric; respiratory syncytial virus; RSV; immunize; ds.  
 XX Unidentified.  
 XX WO200068392-A1.  
 XX 16-NOV-2000.  
 XX 10-MAY-2000; 2000WO-US12582.  
 XX 11-MAY-1999; 99US-0133536.  
 XX (UNII ) UNIV ILLINOIS FOUND.  
 XX Buetow DE, Korban SS, Sandhu J, Krasnyanski SF;  
 WPI; 2001-122707/13.  
 XX Chimeric nucleic acid construct for immunizing animals and humans  
 against respiratory syncytial virus (RSV), comprises a sequence adapted  
 for expression in plants and a RSV protein or peptide coding sequence  
 -  
 XX Disclosure; Fig 10; 67pp; English.  
 XX The present invention relates to a chimeric nucleic acid construct  
 comprising: a nucleotide sequence adapted for protein expression in  
 plants; and a respiratory syncytial virus (RSV) coding sequence  
 encoding an RSV protein or an antigenic protein or peptide of RSV.  
 XX The construct can be used to immunize animals and humans against  
 respiratory syncytial virus. The use of transgenic plants to  
 generate the antigen allows the production of greater amounts of  
 CC antigen.  
 XX Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;



Query Match 92.2%; Score 848.2; DB 22; Length 918;  
Best Local Similarity 95.8%; Pred. No. 2.6e-194;  
Matches 871; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
QY 1 TGCACATGTCACAAAACAGAGGACCAAGCAGCCGCTAGAGACACTAGAAAGACCTGGGA 60
DB 9 TGCACATGTCACAAAACAGAGGACCAAGCAGCCGCTAGAGACACTAGAAAGGACCTGGGA 68
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGCTTTATATAGTTAAATCTTAAATCTGT 120
DB 69 CACTCTCAATCATTTATTTATTCATATCATCGGCTTTATATAGTTAAATCTTAAATCTGT 128
QY 121 AGCACAATCAGATTATTCATTTCTGCAATGATATCTCAACTTCATTTAATTAACG 180
DB 129 AGCACAATCAGATTATTCATTTCTGCAATGATATCTCAACTTCATTTAATTCGCG 188
QY 181 CATCATATTCATAGCTTCGGCAACCAACCAAGTCACACTAAACCTGCAATCATACAAG 240
DB 189 CATCATATTCATAGCTTCGGCAACCAACCAAGTCACACCAACTGCAATCATACAAG 248
QY 241 TGCACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTG 300
DB 249 TGCACAAGCCAGATCAAGAACCAACCCCAACATACCTCACTCAGGATCCTCAGCTTG 308
QY 301 ATTCAGCTTCTCAATCTGCTGAAATTCATCATCAACCAACCACTACCTAGCTTCAAC 360
DB 309 ATTCAGCTTCTCAATCTGCTGAAATTCATCATCAACCAACCACTACCTAGCTTCAAC 368
QY 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAACCAAGTCACAGCTAAAAACCAACAAC 420
DB 369 AACACAGGAGTCAAGTCAAACTGCAACCCCAACCAAGTCACAGCTAAAAACCAACAAC 428
QY 421 AACCCAAACCAACCCAGGAGCCCACTACAAACCAAGCCCAACCAACCAACCAACCA 480
DB 429 AACTCAAAACCAACCCAGGAGCCCACTACAAACCAAGCCCAACCAACCAACCAACCA 488
QY 481 ACCCAATATGTTTCACTTCGAGTGTTCATCTGTTTCACTTGTGAGTATGTCAGCA 540
DB 489 ACCCAATATGTTTCACTTCGAGTGTTCATCTGTTTCACTTGTGAGTATGTCAGCA 548
QY 541 CAATCCAACTGCTGGGCTATCTGCAAGAAATACCAACCAACCAACCAACCAACCAAC 600
DB 549 CAATCCAACTGCTGGGCTATCTGCAAGAAATACCAACCAACCAACCAACCAACCAAC 608
QY 601 CACCAACCAAGCTTCAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660
DB 609 CACTTACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTCA 668
QY 661 AACCACTAACCAAGGAGTACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 720
DB 669 AACCACTAACCAAGGAGTACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 728
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 780
DB 729 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 788
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
DB 789 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 848
QY 841 AGTCTCCACATCCGAGGAGCCATCAGACCTCTCTCCGAGGCAATCTAAGCCCTTCTCA 900
DB 849 AGTCTCTACACATCCGAGGAGCCATCAGACCTCTCTCCGAGGCAATCTAAGCCCTTCTCA 908
QY 901 GTAGTTATT 909
DB 909 GTAGTTACT 917
```

RESULT 11

AAS20145

ID AAS20145 standard; DNA; 897 BP.

XX

AC AAS20145;  
XX 09-APR-2002 (first entry)  
XX Respiratory syncytial virus G protein DNA.

XX RSV; ds; G protein; heavily glycosylated protein; antianemic; antiviral;  
XX vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
XX antiviral chemotherapeutic compound; humoral response;  
XX cellular immune response; HIV; paediatric respiratory disease;  
XX globin gene transfer; sickle cell disease; beta-thalassaemia;  
XX human immunodeficiency virus infection; HIV.

XX Human respiratory syncytial virus.

XX Key Location/Qualifiers  
CDS 1..897  
/\*tag= a  
/product= "G protein"

XX WO200192548-A2.

XX 06-DEC-2001.

XX 22-MAY-2001; 2001WO-US16610.

XX 01-JUN-2000; 2000US-208701P.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Portner A, Takimoto T;

XX WPI; 2002-130534/17.

XX P-PSDB; AAU74676.

XX Recombinant Sendai virus useful in vaccines to protect infection by  
PT paramyxoviruses, comprises exogenous nucleic acid encoding  
PT paramyxovirus protein or its antigenic fragment -

XX Disclosure; Page 47; 57pp; English.

XX The invention relates to a recombinant Sendai virus comprising an  
CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
CC antigenic fragment. The virus may be administered in combination  
CC with an antiviral chemotherapeutic compound. Two or more viruses  
CC expressing different PMV proteins may be co-administered. Compositions  
CC comprising the virus are useful for eliciting a humoral and/or  
CC cellular immune response to a PMV in a mammal, particularly a human.

CC Further a recombinant Sendai virus comprising an exogenous nucleic acid  
CC encoding a second PMV protein is also administered and priming and/or  
CC boosting humoral or cellular immune response comprises administering  
CC one or more of a recombinant or isolated PMV protein or its antigenic  
CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
CC vector encoding a PMV protein. The recombinant virus is useful as an  
CC effective vaccine against HIV or RSV (the major causes of paediatric  
CC respiratory disease) and also to express any gene of

CC interest in target cells, providing a positive medical impact on  
CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
CC into stem cells effects a cure for sickle cell disease or beta-  
CC thalassaemia. The recombinant virus may also prove effective in  
CC conferring immunity to human immunodeficiency virus (HIV) infection.  
CC The Sendai virus replicates at level that is high enough to  
CC induce sufficient immunity, but does not cause any harm to human  
CC recipient. The present sequence encodes a respiratory syncytial  
CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
CC suitable for expression by the recombinant virus of the invention.

XX Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other;

XX Query Match 91.1%; Score 837.8; DB 24; Length 897;

XX Best Local Similarity 95.9%; Pred. No. 8e-192;

XX Matches 860; Conservative 0; Mismatches 37; Indels 0; Gaps 0;



QY 481 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTTACCTGCGAGCATATGCGACAA 540  
Db 10056 ACCCAATATGATTTTCACTTGGAGTGTAACTTTGTTACCTGCGAGCATATGCGACAA 9997  
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAAACCAAGAAAGAAAC 600  
Db 9996 CAATCCAACTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGAAAGAAAC 9937  
QY 601 CACACCAAGCTTACAAAACCAAACTTCAAGACAAACCAAAAGAAATCTCAAACTCA 660  
Db 9936 CACTTACCAAGCTTACAAAACCAAACTTCAAGACAAACCAAAAGAAATCTCAAACTCA 9877  
QY 661 AACCACTTAAACCAAGAAAGTACCAACCAAGGCGCCACAGAGGCGCCACCATCAAC 720  
Db 9876 AACCACTTAAACCAAGAAAGTACCAACCAAGGCGCCACAGAGGCGCCACCATCAAC 9817  
QY 721 CACCAAAACCAACATCACTACTGCTCACCACCAACCAACCAAGGAAATCCAAACT 780  
Db 9816 CACCAAAACCAACATCACTACTGCTCACCACCAACCAACCAAGGAAATCCAGACT 9757  
QY 781 CACAAGTCAATGAAACCTTCACTCACTCACTTCTTCTCGAAGGCAATCTAAGCCCTTCTCA 840  
Db 9756 CACAAGTCAATGAAACCTTCACTCACTCACTTCTTCTCGAAGGCAATCTAAGCCCTTCTCA 9697  
QY 841 AGTCTCACACATCCGAGCACCACCATCACACACCTCATCTCCACCAACCAACAGCCCA 900  
Db 9696 AGTCTCACACATCCGAGCACCACCATCACACACCTCATCTCCACCAACCAACAGCCCA 9637  
QY 901 GTAGTTATTAAAAAAA 919  
Db 9636 GTAGTTATTAAAAACATA 9618

## RESULT 13

AAK08422

ID AAK08422 standard; cDNA; 715 BP.

XX AC AAX08422;

XX DT 28-JUN-1999 (first entry)

XX G protein gene fragment of respiratory syncytial virus.  
XX G protein: respiratory syncytial virus; RSV; recombinant vector;  
XX vaccine; immune response; immunogenicity; tPA; antibody;  
XX tissue plasminogen activator; ss.

XX Respiratory syncytial virus (RSV).  
XX Key Location/Qualifiers  
XX CDS 1..702  
XX /tag= a  
XX /product= "Secreted G protein"

XX WO9904010-A1.

XX 28-JAN-1999.

XX 16-JUL-1998; 98WO-CA00697.

XX 18-JUL-1997; 97US-0896442.

XX (CONN-) CONNAUGHT LAB LTD.

XX Klein MH, Li X, Sambhara S;

XX WPI: 1999-132254/11.

XX P-PSDB; AAW96314.

XX Immunogenic composition for generating antibodies against  
XX respiratory syncytial virus - comprises non-replicating vector  
XX containing the protein G sequence, useful in protective vaccines and  
XX to raise antibodies for diagnosis

XX Claim 8; Figure 3; 67pp; English.  
XX

CC The respiratory syncytial virus (RSV) G protein can be used in  
CC vaccines by inserting the G protein gene into a non-replicating  
CC vector. The G protein is placed under the control of alternative  
CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G protein's  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.

XX Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;  
SQ

Query Match 77.7%; Score 715; DB 20; Length 715;  
Best Local Similarity 100.0%; Pred. No. 2.5e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 CACAAAGTCACACTAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 265  
Db 1 CACAAAGTCACACTAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60

QY 266 ACCCCAACTACTCTACTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAACTCTGTCTGAA 325  
Db 61 ACCCCAACTACTCTACTCAGGATCTCTAGCTTGGAAATCAGCTTCTCCAACTCTGTCTGAA 120

QY 326 ATTACATCAAAACACACACCATATAGCTTCAACACACCGAGGATCAAGTCAAACTCG 385  
Db 121 ATTACATCAAAACACACCATATAGCTTCAACACACCGAGGATCAAGTCAAACTCG 180

QY 386 CAAACCAACAGTCAAGACTTAAACACACACCAACCAACCAACCAACCAACCAACCAACCA 445  
Db 181 CAAACCAACAGTCAAGACTTAAACACACACCAACCAACCAACCAACCAACCAACCAACCA 240

QY 446 ACTACAAAACACCGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505  
Db 241 ACTACAAAACACCGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300

QY 506 GTGTTTAACTTTGTACCTTGCAGCATATGCGAGCAACATCCACCTGTGGGTATCTGC 565  
Db 301 GTGTTTAACTTTGTACCTTGCAGCATATGCGAGCAACATCCACCTGTGGGTATCTGC 360

QY 566 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625  
Db 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420

QY 626 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685  
Db 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480

QY 686 ACCACCAAGCCACAGAGGCGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 745  
Db 481 ACCACCAAGCCACAGAGGCGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540

QY 746 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805  
Db 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600

QY 806 TCAACCTCTCCGAGGCGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCACA 865  
Db 601 TCAACCTCTCCGAGGCGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCACA 660

QY 866 TCAACCTCTCTATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920  
Db 661 TCAACCTCTCTATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715

RESULT 14

AAV38298  
ID AAV38298 standard; DNA; 696 BP.  
XX  
AC AAV38298;  
XX  
DT 26-OCT-1998 (first entry)  
XX  
XX Respiratory syncytial virus glycoprotein G gene portion.  
DE RSV; glycoprotein G; pneumococcal surface protein A; PspA;  
XX infection; Streptococcus pneumoniae; sepsis; otitis media;  
KW meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;  
KW ss.  
XX  
XX Respiratory syncytial virus.  
XX  
XX WO9824927-A1.  
XX  
XX 11-JUN-1998.  
XX  
XX 04-DEC-1997; 97WO-US22847.  
XX  
XX 04-DEC-1996; 96US-0759505.  
XX  
XX (UYAL-) UNIV ALABAMA.  
XX  
XX Briles DE, Curriel DT, McDaniel LS;  
XX  
XX WPI; 1998-333343/29.  
XX  
XX Plasmid containing pneumococcal epitope for expression in eukaryotic  
PT cells - useful for eliciting immunological response to pneumococcal  
PT infection or sepsis  
XX  
XX Example 1; Fig 1C; 47pp; English.  
XX  
XX This is a portion of the respiratory syncytial virus glycoprotein G  
CC (RSVG) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)  
CC to create plasmid pGT4. This plasmid contains a human  
CC cytomegalovirus immediate early promoter and the RSVG gene portion  
CC such that when an in-frame fusion is made, the resultant fusion  
CC protein may be transported to, and anchored in, a mammalian cell  
CC membrane where it can be exposed to the host immune system.  
CC Insertion of pneumococcal surface protein A (PspA) coding sequence  
CC created plasmid pKSD2601. Intramuscular immunisation of BALB/c  
CC mice with pKSD2601 induced protection against an otherwise lethal  
CC challenge with a capsular type 3 pneumococcus. A claimed plasmid  
CC for expression of pneumococcal epitope DNA in eukaryotic cells  
CC includes a promoter for driving expression in a eukaryotic cell  
CC (e.g. HCMV-IE). DNA encoding a leader sequence (e.g. of RSVG)  
CC and DNA encoding a pneumococcal epitope such as PspA. The invention  
CC also provides a vaccine comprising the plasmid and a suitable  
CC carrier or diluent, and optionally one or more cytokines or DNA  
CC encoding them, or a bacterial delivery system. The vaccine is used  
CC to elicit an immunological response in a host, including humans,  
CC susceptible to pneumococcal infection or sepsis. The plasmid can  
CC also be used to express a pneumococcal epitope of interest in vitro.  
XX  
SQ Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;  
Query Match 69.3%; Score 637.4; DB 19; Length 696;  
Best Local Similarity 96.1%; Pred. No. 1.1e-143;  
Matches 664; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
8 ATGTCGAAAACAAAGGACCAACCGCTAGACACTAGAAAGACCTGGGACACTCTC 67  
1 ATGTCGAAAACAAAGGACCAACCGCTAGACACTAGAAAGACCTGGGACACTCTC 60  
68 AATCATTATTATTCATATCATCGGCTTATATAGTTAACTTAAATCTGTAGACAA 127  
61 AATCATTATTATTCATATCATCGGCTTATATAGTTAACTTAAATCTGTAGACAA 120  
128 ATCATTATTCATTCCTGGCAATGATAATCTCACTTATATATAGCCATCATA 187

Db 121 ATCATTATTCATTCCTGGCAATGATAATCTCACTTATATATAGCCATCATA 180  
Qy 188 TTCATAGCCTCGGCAAAACCAAGTCACTAACTCACTCAATCATCAAGATGCAACA 247  
Db 181 TTCATAGCCTCGGCAAAACCAAGTCACTAACTCACTCAATCATCAAGATGCAACA 240  
Qy 248 AGCCAGATCAAGAACACACACCCCAATACCTCACTCAGGATCTCAGTTGGAATCAGC 307  
Db 241 AGCCAGATCAAGAACACACACCCCAATACCTCACTCAGGATCTCAGTTGGAATCAGT 300  
Qy 308 TTCTCCAAATCTGTCTGAAATTCATATCAACCAACCACTACTAGTTTCAACACCA 367  
Db 301 CCCTCTAATCCGCTGAAATTCATATCAAAATCAACCACTACTAGTTTCAACACCA 360  
Qy 368 GGAGTCAAGTCAAAACCTGCAACCCCAACCACTCAAGTCAAGTCAAAACCAACCAAC 426  
Db 361 GGAGTCAAGTCAAAACCTGCAATCCACCACTCAAGTCAAGTCAAAACCAACCAAC 420  
Qy 427 AACACAAACCCAGCAGCCCACTCAAAACCAACCAACCAACCAACCAACCAACCA 486  
Db 421 AACACAAACCCAGCAGCCCACTCAAAACCAACCAACCAACCAACCAACCAACCA 480  
Qy 487 TAATGATTTTCACTTCGAGTGTGTTAACTTTGTAACCTGCGAGCATATGCGAATCC 546  
Db 481 TAATGATTTTCACTTTGAGTGTGTTAACTTTGTAACCTGCGAGCATATGCGAATCC 540  
Qy 547 AACCTGCTGGGCTATCTGCAAAAGATAATCCAAACCAAAACCAAGGAAAGAAACCCAC 606  
Db 541 AACCTGCTGGGCTATCTGCAAAAGATAATCCAAACCAAAACCAAGGAAAGAAACCCAC 600  
Qy 607 CAAGCCTACAAAACCAACCACTTCAAGCAACCAAAACCAAGATCTCAAACCTCAAACCCAC 666  
Db 601 CAAGCCTACAAAACCAACCACTTCAAGCAACCAAAACCAAGATCTCAAACCTCAAACCCAC 660  
Qy 667 TAAACCAAGGAGTACCCACCAACCCAGGCC 697  
Db 661 TAAATCAAGGAGTACCCACCAACCCAGGCC 691

RESULT 15  
AAV17552  
ID AAV17552 standard; cDNA; 15225 BP.  
XX AAV17552;  
AC AAV17552;  
XX  
DT 20-JUL-1998 (first entry)  
XX  
DE Respiratory syncytial virus genome.  
XX  
KW RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.  
XX  
OS Human respiratory syncytial virus B-1.  
XX  
PN WO9802530-A1.  
XX  
PD 22-JAN-1998.  
XX  
XX 15-JUL-1997; 97WO-US12269.  
XX  
XX 23-MAY-1997; 97US-0047634.  
PR 15-JUL-1996; 96US-0021773.  
PR 09-MAY-1997; 97US-0046141.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;  
PI Whitehead SS;  
XX  
XX WPI; 1998-110579/10.  
XX  
XX Attenuated respiratory syncytial virus vaccines - useful to protect  
PT individuals against RSV infection

XX PS

Example 4; Page 195-202; 238pp; English.

This is the complete nucleotide sequence of the wild-type B-1 respiratory syncytial virus (RSV). The genome is negative-sense; the 5'-3' positive-sense sequence of D46 is provided in AAV17553. A novel infectious recombinant RSV comprises a RSV genome or antigenome, a major nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a large polymerase protein (L), and a RNA polymerase elongation factor, where the recombinant RSV has at least two attenuating mutations, one of the mutations specifying a temperature-sensitive (ts) substitution at amino acid Phe521, Gln331, Met1169 or Tyr1321 in the RSV polymerase gene or a ts nucleotide substitution in the gene-start sequence of gene M2. Also claimed are: (i) an isolated infectious RSV particle which comprises a recombinant RSV (antigenome, N, P, and L proteins, a RNA polymerase elongation factor, where the (antigenome is modified (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or a cis-acting regulatory sequence; and (ii) by a termination codon introduced within a selected gene, or by a change in sequence, position or presence of a GS or GE transcription signal relative to the selected gene; (2) an expression vector; and (3) an RSV strain selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452), cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or B-1 cp52/2b5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated attenuated recombinant RSV and RSV particles are used in a vaccine to stimulate the immune system of an individual to induce protection against RSV. The expression vector of (2) is used for the production of infectious attenuated RSV particles.

Sequence 15225 BP; 5904 A; 2710 C; 2399 G; 4212 T; 0 other;

Query Match 48.0%; Score 441.4; DB 19; Length 15225;  
Best Local Similarity 68.1%; Pred. No. 3.8e-96;  
Matches 629; Conservative 0; Mismatches 291; Indels 3; Gaps 1;  
1 TGCACACATGTCTCAAAAAACAAGGACCAACGCCCGCTAAGACACTAGAAAAGACCTGGG 60  
4683 TGCACACATGTCTCAAAACAACAGATCAACGACTGCCAGGACTCTAGAAAAGACCTGGG 4742

Qy	1	TGCAACATGTCCAAAAACAAGGACCAACGCGCCGCTTAAGACACTAGAAAAGACCTGGGA	60
Db	4683	TGCAACCATGTCCAAACAACAAGAAATCAACGCACTGCCAGGACTCTAGAAAAGACCTGGGA	4742

Qy	61	CAC	TCT	CAAT	CA	TTT	TAT	TAT	TAT	CA	TAT	CA	TCT	CGG	CTT	TAT	TAT	AT	AGT	TAA	AT	CTGT	120
D <sub>b</sub>	4743	TACT	TCT	CAAT	CA	TCT	TAAT	TGT	AA	TAT	TCT	CTT	CTT	TAT	AT	CA	GAT	TAA	AT	TAA	AT	CTAT	4802

Qy	121	AGCACAAATCACTATTCCATTCTGGCAATGATTAATCTCAACTTCATTATTAATTCAGC	180
Dβ	4803	AGCACAAATAGCACTATCAGTTCTGGCAATGATTAATCTCAACTCTCTCATATTTGCGC	4862

Qy	181	CATCATATTTCATAGCCTCGGCAACCCAAAGTCACACTAACAACTGCATCATACAGA	240
Db <td>4863</td> <td>CATAATTCATCATCTCTGCGCAATCACAAAGTTTACCTAATCAACGGTCACTGCTCAAC</td> <td>4922</td>	4863	CATAATTCATCATCTCTGCGCAATCACAAAGTTTACCTAATCAACGGTCACTGCTCAAC	4922

QY 241 TGCAACAGCCAGATCAAGAAACACAAACCCCAACATACCTCACTCAGAGTCCCTCAGCTTG 300

Dh 4923 AATAAAAAACCACTGAAAAAACAATCAACCTACTCTTCTTCAAGATCCACACAGAAAG 4982

**OY**

301 AATCAGCTTCTCCAATTGTCGTGAATAATACATCACAAACCACCACTACTAGCTTCAAC 360  
|||||  
**DH**

4983 GATTATGCATCATCTTTGGTAACAACCTACACACCGCCATGCCAATGCCATCGCGCGGCTGAC 5043  
|||||

QY 361 AACACAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTATATAAACACCAACAAC 420

421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACAAACGCCAAACCAACCAACAA 480

Q2 481 ACCCAATAATGATTTTCACCTTCGAAGCTTTTAACTTTGTACCCCTGCACCATATGCAGCA 540

---

Qy	541	CAATCCAACTGCTGGCTATCTTGGCAAGAATACAAACAAAAAACCGAGGAAGAAAC	600
Db	5223	CAATCAACTTTGGCAATTCCTCTGTGTAAACAATATACCAAGCAACAACCAAGAGAAAC	5282
Qy	601	CACCACCGAGCTTACAAAAAACAACCACTTCAA--GACACCAAAAAAGATCTCAAC	657
Db	5283	AACCATCAACCCACAACAACAACAACCAACCAACCAAAACCAAAAGAGACCAAAAA	5342
Qy	658	TCAAAACCATTAACCAAGGAGTACCCACCAAGCCCAAGAGAGCCAAACCATCAA	717
Db	5343	ACCAGCCAAAACGACGAAAAAGAACTACCAACCAACCAAAAAACCAACCTCAC	5402
Qy	718	CACCAACCAAAAAACATCAACAATCACTGTCTCAACAACAACCAACAGGAAATCCAAA	777
Db	5403	GAGCCACAGAAAGAGACACCAAGCACTTCACAATCCACTGTGCTCGACACCAACATAGA	5462
Qy	778	ACTCAAGTCAATGGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTC	837
Db	5463	ACACACATCCAAAGCAATCCCTTCACTCAACCCCGGAAAAACACCCCAACTCCAC	5522
Qy	838	TCAAGTCTCCAACAACATCCGAGACCCCATCAACAACCTCATCTTCCACCAACAACAACG	897
Db	5523	ACAAACCCACACAGCATCCGAGCCCTCTACATCAATTCACCCAAAATACCAATCACA	5582
Qy	898	CCAGTAGTTATTAATAAAAAA 920	
Db	5583	TGCTTAGTTATTCAAAAACTACA 5605	

Search completed: October 29, 2003, 22:09:33  
Job time : 316.982 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:51:32 ; Search time 3762.15 Seconds  
(without alignments)  
10004.074 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcgaacatgccccaaaaca.....gtagttattataaaaaaaaaa 920

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htg\_hum.\*  
40: em\_htg\_mus.\*  
41: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	920	100.0	920	6	AI6257	AI6257 Synthetic R
2	920	100.0	920	6	AI6258	AI6258 Synthetic R
3	920	100.0	920	6	AR080406	AR080406 Sequence
4	920	100.0	920	6	AR092530	AR092530 Sequence
5	920	100.0	920	6	AR122885	AR122885 Sequence
6	920	100.0	920	6	AR123540	AR123540 Sequence
7	920	100.0	920	6	AR148357	AR148357 Sequence
8	920	100.0	920	6	BD081933	BD081933 Nucleic a
9	911.8	99.1	923	14	HRSNNG	XI7085 Human respi
10	909.2	98.8	922	14	HRSVGL16	Z33429 Human respi
11	901.6	98.0	917	14	RSHGLYG	M17212 Human respi
12	894	97.2	894	6	AR080424	AR080424 Sequence
13	894	97.2	894	6	AR092548	AR092548 Sequence
14	853.4	92.8	8510	14	RSHICE	M11486 Human respi
15	853.4	92.8	15222	6	AR093219	AR093219 Sequence
16	853.4	92.8	15222	6	BD091237	BD091237 RNase L a
17	853.4	92.8	15222	14	HRU50362	U50362 Human respi
18	853.4	92.8	15222	14	HRU50363	U50363 Human respi
19	853.4	92.8	15222	14	HRU63644	U63644 Human respi
20	853.4	92.8	15222	14	RSHSEQ	M74568 Human respi
21	853.4	92.8	15223	6	AR089137	AR089137 Sequence
22	853.4	92.8	15223	14	AF035006	AF035006 Human res
23	848.2	92.2	914	14	PARSENVG	X03149 Respiratory
24	848.2	92.2	918	6	AX047075	AX047075 Sequence
25	837.8	91.1	897	6	AX339011	AX339011 Sequence
26	837.2	91.0	15190	14	HRU39662	U39662 Human respi
27	835.6	90.8	15191	14	RSU39661	U39661 Respiratory
28	820.4	89.2	918	14	AF065406	AF065406 Human res
29	814	88.5	918	14	AF065405	AF065405 Human res
30	811.6	88.2	922	14	HRSCVGL1	Z33410 Human respi
31	811.6	88.2	922	14	HRSCVGA	Z33454 Human respi
32	810	88.0	922	14	HRSCVG3	Z33412 Human respi
33	810	88.0	922	14	HRSCVGC	Z33456 Human respi
34	810	88.0	922	14	HRSVGL11	Z33424 Human respi
35	810	88.0	922	14	HRSVGL12	Z33425 Human respi
36	808.4	87.9	922	14	HRSVGGLY1	Z33493 Human respi
37	808.4	87.9	922	14	HRSVGGLY2	Z33494 Human respi
38	807.2	87.7	924	14	HRSVGL14	Z33427 Human respi
39	806.8	87.7	922	14	HRSVGL2	Z33415 Human respi
40	806.8	87.7	922	14	HRSVGL9	Z33422 Human respi
41	805.2	87.5	922	14	AF065256	AF065256 Human res
42	805.2	87.5	922	14	HRSVGL10	Z33423 Human respi
43	805.2	87.5	922	14	HRSVGL13	Z33426 Human respi
44	805.2	87.5	922	14	HRSVGL3	Z33416 Human respi
45	803.6	87.3	922	14	AF065255	AF065255 Human res

# ALIGNMENTS

RESULT 1  
AL6257  
LOCUS AL6257 Synthetic RSV G gene (seq ID No: 7).  
DEFINITION AL6257  
ACCESSION AL6257  
VERSION AL6257.1 GI:640933  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS CHIMERIC IMMUNOGENS  
TITLE Patent: WO 9314207-A 7 22-JUL-1993;  
JOURNAL Location/Qualifiers  
FEATURES linear PAT 03-OCT-1994

```
source 1. 920
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
8. .904
/gene="RSV G"
8. .904
/gene="RSV G"
/codon_start=1
/transl_table=11
/protein_id="CAA01273.1"
/db_xref="GI:640934"
/translation="MSKNDQRTAKTLEKWTWDTLNHLLFISGLYKMLKSVQIOTLS
ILAMILISLIITAIPTFASANKVLTITAIQDATSQIKNTPTPYLTQDQULGISFS
NLSEITSLITLITLTPGVKSNLQPTTKTKNTTTTTQPSKPTTKQRQKNPKNPQ
NDFPEVFNVPSCISNNPTCWAICRIIPNKKPGKTKTKTKPTFKTKKDLKPKQ
TKKPEVPTTKPTTEPTINTTKNTITTTTTLLTNNTGNPKLTSQMETFHTSSEGNLSP
SQVSTTSBHPSPSPSPNTTQ"
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3e-192;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCACACATGTCCTCAAAACAAAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
DB 1 TGCACACATGTCCTCAAAACAAAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
QY 61 CACTCTCAATCATTTATTATTCATATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120
DB 61 CACTCTCAATCATTTATTATTCATATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120
QY 121 AGCACAATFCAATATCAATCTCTGGCAATGATATCTCAATCTCACTTATATTAATACAGC 180
DB 121 AGCACAATFCAATATCAATCTCTGGCAATGATATCTCAATCTCACTTATATTAATACAGC 180
QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGCACTGCAATCATACAGA 240
DB 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGCACTGCAATCATACAGA 240
QY 241 TGCACAAGCCAGATCAAGAACCAACCAACCAATACCTCTCAGGAGTCTTCAGCTGG 300
DB 241 TGCACAAGCCAGATCAAGAACCAACCAACCAATACCTCTCAGGAGTCTTCAGCTGG 300
QY 301 AATCAGCTTCTCCATCTGTCTGAAATTAATCAATCAAAACCAACCAATCTAGCTTCAAC 360
DB 301 AATCAGCTTCTCCATCTGTCTGAAATTAATCAATCAAAACCAACCAATCTAGCTTCAAC 360
QY 361 AACACACAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACACAAAC 420
DB 361 AACACACAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACACAAAC 420
QY 421 AACCCCAACACAAACCCAGACCCCACTACAAAACCAAGCCCAACCAACCAACCAAC 480
DB 421 AACCCCAACACAAACCCAGACCCCACTACAAAACCAAGCCCAACCAACCAACCAAC 480
QY 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGCAGCATATGAGCAA 540
DB 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGTACCTCGCAGCATATGAGCAA 540
QY 541 CAATCCACCTGCTGGGCTATCTGCAAGATATCAACCAACCAACCAACCAACCAACCAAC 600
DB 541 CAATCCACCTGCTGGGCTATCTGCAAGATATCAACCAACCAACCAACCAACCAACCAAC 600
QY 601 CACCAACCAAGCTTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660
DB 601 CACCAACCAAGCTTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660
QY 661 AACCACTAAACCAAGGAGTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 720
DB 661 AACCACTAAACCAAGGAGTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 720
```

```
QY 721 CACCACAAACACATCACAACCTACACTGCTCACCAACCAACCAACCAAGAAATCCAAACT 780
DB 721 CACCACAAACACATCACAACCTACACTGCTCACCAACCAACCAACCAAGAAATCCAAACT 780
QY 781 CACAAGTCAAAATGGAACCTTCCACTCAACTCAACTCTCCGAAGGCAATCTTAAGCCCTTCTCA 840
DB 781 CACAAGTCAAAATGGAACCTTCCACTCAACTCTCCGAAGGCAATCTTAAGCCCTTCTCA 840
QY 841 AGTCTCCACAAATCCGAGCACCCTCATCAACCCCTCATCTCTCCACCCACACACACGCCA 900
DB 841 AGTCTCCACAAATCCGAGCACCCTCATCAACCCCTCATCTCTCCACCCACACACACGCCA 900
QY 901 GTAGTTATTAATAAAAAA 920
DB 901 GTAGTTATTAATAAAAAA 920
RESULT 2
LOCUS A16258 920 bp DNA linear PAT 03-OCT-1994
DEFINITION Synthetic RSV G gene (seq ID No: 7).
ACCESSION A16258
VERSION A16258.1 GI:640935
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 920)
AUTHORS .
TITLE CHIMERIC IMMUNOGENS
JOURNAL Patent: WO 9314207-A 27 22-JUL-1993;
FEATURES
location/Qualifiers
1. .920
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 155 a 95 c 290 g 380 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3e-192;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCACACATGTCCTCAAAACAAAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
DB 920 TGCACACATGTCCTCAAAACAAAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 861
QY 61 CACTCTCAATCATTTATTATTCATATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120
DB 860 CACTCTCAATCATTTATTATTCATATCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 801
QY 121 AGCACAATFCAATATTCATTTCTGGCAATGATATCTCAATCTCACTTATATTAATACAGC 180
DB 800 AGCACAATFCAATATTCATTTCTGGCAATGATATCTCAATCTCACTTATATTAATACAGC 741
QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGCACTGCAATCATACAGA 240
DB 740 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAAGTCACACTGCAATCATACAGA 681
QY 241 TGCACAAGCCAGATCAAGAACCAACCAACCAATACCTCTCAGGAGTCTTCAGCTGG 300
DB 680 TGCACAAGCCAGATCAAGAACCAACCAACCAATACCTCTCAGGAGTCTTCAGCTGG 621
QY 301 AATCAGCTTCTCCATCTGTCTGAAATTAATCAATCAAAACCAACCAACCAACCTAGCTTCAAC 360
DB 620 AATCAGCTTCTCCATCTGTCTGAAATTAATCAATCAAAACCAACCAACCAACCTAGCTTCAAC 561
QY 361 AACACACAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACACAAAC 420
DB 560 AACACACAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGCTTAAACACACAAAC 501
QY 421 AACCCCAACACAAACCCAGACCCCACTACAAAACCAAGCCCAACCAACCAACCAAC 480
```

```

Db      500 AACCCAAACACAAACCCAGCAGCCCACTACAAAACAAACGCCCCAAACAAACACCAACCAAA 441
Qy      481 ACCCAATATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTCGCAGCATATGCGCAA 540
Db      440 ACCCAATATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTCGCAGCATATGCGCAA 381
Qy      541 CAATCAACCTGCTGGGCTATCTGCAAAAGATATACAAACAAACAAACCAAGGAAAGAAAC 600
Db      380 CAATCAACCTGCTGGGCTATCTGCAAAAGATATACAAACAAACAAACCAAGGAAAGAAAC 321
Qy      601 CACGACCAAGCTTACAAAACCAACCTTCAAGACAAACCAACCAACCAACCAACCAACCAAC 660
Db      320 CACCAACCAAGCTTACAAAACCAACCTTCAAGACAAACCAACCAACCAACCAACCAACCAAC 261
Qy      661 AACCACTAAACCAAGGAGTACCCACCAAGGAGCCACAGAGAGCCCAACCATCAACAC 720
Db      260 AACCACTAAACCAAGGAGTACCCACCAAGGAGCCCAACCATCAACAC 201
Qy      721 CACCAAAACAAACATCAACACTTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAAC 780
Db      200 CACCAAAACAAACATCAACACTTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAAC 141
Qy      781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Db      140 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 81
Qy      841 AGTCTCCAAACATCCGAGCAGCCATCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 900
Db      80 AGTCTCCAAACATCCGAGCAGCCATCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 21
Qy      901 GTAGTTATTAATAAAAAA 920
Db      20 GTAGTTATTAATAAAAAA 1

RESULT 3
AR080406
LOCUS      AR080406      920 bp      DNA      linear      PAT 31-AUG-2000
DEFINITION Sequence 7 from patent US 5968776.
ACCESSION AR080406
VERSION    AR080406.1 GI:10007141
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 920)
AUTHORS     Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE       Multimeric hybrid gene encoding a chimeric protein which confers
            protection against parainfluenza virus and respiratory syncytial
            virus
JOURNAL     Patent: US 5968776-A 7 19-OCT-1999;
FEATURES    Location/Qualifiers
            1..920
            source          380 a 290 c 95 g 155 t

Query Match      100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3e-192;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TGCACCAATGTCACAAACCAAGGAGCCACCGCTAAGCAGCAGTACGAAAGACCTGGGA 60
Db      1 TGCACCAATGTCACAAACCAAGGAGCCACCGCTAAGCAGCAGTACGAAAGACCTGGGA 60
Qy      61 CACTCTCAATCATTTATTTATCATATCATCGGGCTATATAAGTTAAATCTTAATCTGT 120
Db      61 CACTCTCAATCATTTATTTATCATATCATCGGGCTATATAAGTTAAATCTTAATCTGT 120
Qy      121 AGCACAAATCAGATTATTCATTTCTGGCAATGATAATCTCAACTTCATTATAATTACAGC 180
Db      121 AGCACAAATCAGATTATTCATTTCTGGCAATGATAATCTCAACTTCATTATAATTACAGC 180

```

```

Qy      181 CATCATATTTAGCTCGGCACAAACCAAGTCACACTAACACTGCAATCATACAAGA 240
Db      181 CATCATATTTAGCTCGGCACAAACCAAGTCACACTAACACTGCAATCATACAAGA 240
Qy      241 TGCACCAAGCCAGATCAAGAACACAAACCCCAATACCTCTCAGGATCTCTCAGCTTGG 300
Db      241 TGCACCAAGCCAGATCAAGAACACAAACCCCAATACCTCTCAGGATCTCTCAGCTTGG 300
Qy      301 AATCAGCTTCTCAATCTGTCTGAAATTTACATCAAAACCAACCAACCACTACTAGCTTCAAC 360
Db      301 AATCAGCTTCTCAATCTGTCTGAAATTTACATCAAAACCAACCAACCACTACTAGCTTCAAC 360
Qy      361 AACCAACCAAGGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
Db      361 AACCAACCAAGGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
Qy      421 AACCAACCAACCAAGGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAG 480
Db      421 AACCAACCAACCAAGGAGTCAAGTCAACCTGCAACCAACCAAGTCAAGTCAAGTCAAGTCAAG 480
Qy      481 ACCCAATATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTCGCAGCATATGCGCAA 540
Db      481 ACCCAATATGATTTTCACTTCGAAAGTGTTTAACTTTGTACCTCGCAGCATATGCGCAA 540
Qy      541 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGGAGGAGGAGGAGGAGGAGG 600
Db      541 CAATCCCAACCTGCTGGGCTATCTGCAAAAGATACCAAAACCAAGGAGGAGGAGGAGGAGGAGG 600
Qy      601 CACCAACCAAGCTTCAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAACCTCA 660
Db      601 CACCAACCAAGCTTCAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAACCTCA 660
Qy      661 AACCACTAAACCAAGGAGTACCCACCAAGGAGCCACAGAGGAGCCACCATCAACAC 720
Db      661 AACCACTAAACCAAGGAGTACCCACCAAGGAGCCACAGAGGAGCCACCATCAACAC 720
Qy      721 CACCAACCAACCAACATCAACACTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAAC 780
Db      721 CACCAACCAACCAACATCAACACTACACTGCTCAACCAACCAACCAACCAACCAACCAACCAAC 780
Qy      781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Db      781 CACAAGTCAATGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840
Qy      841 AGTCTCCAAACATCCGAGCAGCCATCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 900
Db      841 AGTCTCCAAACATCCGAGCAGCCATCAACCTTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 900
Qy      901 GTAGTTATTAATAAAAAA 920
Db      901 GTAGTTATTAATAAAAAA 920

RESULT 4
AR092530
LOCUS      AR092530      920 bp      DNA      linear      PAT 08-SEP-2000
DEFINITION Sequence 7 from patent US 5998169.
ACCESSION AR092530
VERSION    AR092530.1 GI:10019284
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 920)
AUTHORS     Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE       Multimeric hybrid gene encoding a chimeric protein which confers
            protection against parainfluenza virus and respiratory syncytial
            virus
JOURNAL     Patent: US 5998169-A 7 07-DEC-1999;
FEATURES    Location/Qualifiers
            1..920
            source          380 a 290 c 95 g 155 t

```





781 CACAAGTCAATGGAACCTTTCCACTCACTCCTCCGAGGGAATCAAGCCCTTCTCA 840  
Db CACAAGTCAATGGAACCTTTCCACTCACTCCTCCGAGGGAATCAAGCCCTTCTCA 840  
781 CACAAGTCAATGGAACCTTTCCACTCACTCCTCCGAGGGAATCAAGCCCTTCTCA 840  
841 AGTCTCCACAAATCCGAGCACCACATCAACCCCTCATCTCCACCCCAACACAGCCCA 900  
Db AGTCTCCACAAATCCGAGCACCACATCAACCCCTCATCTCCACCCCAACACAGCCCA 900  
901 GTAGTTATTAAAAA 920  
Db GTAGTTATTAAAAA 920

RESULT 6  
AR123540  
LOCUS AR123540 920 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 7 from patent US 6171783.  
ACCESSION AR123540  
VERSION AR123540.1 GI:14108901  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
TITLE Infection detection method using chimeric protein  
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;  
FEATURES Location/Qualifiers  
source  
1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACAACTGTCCAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
Db 1 TGCACAACTGTCCAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
QY 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
Db 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
QY 121 AGCACAATCAATATTCATATTCGCAATGATAATCTCACTTCACTTATTAATACAGC 180  
Db 121 AGCACAATCAATATTCATATTCGCAATGATAATCTCACTTCACTTATTAATACAGC 180  
QY 181 CATCATATTCATAGCTTCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
Db 181 CATCATATTCATAGCTTCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
QY 241 TGCACAAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 241 TGCACAAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
QY 301 ATTCAGCTTCTCAATCTGTCTGAAATTAATCAATCAACCAACCAACCAACCAACCA 360  
Db 301 ATTCAGCTTCTCAATCTGTCTGAAATTAATCAATCAACCAACCAACCAACCAACCA 360  
QY 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420  
Db 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420  
QY 421 AACCCAAACACAAACCCAGCAGCCACTACAAAACAAAGCCCAACCAACCAACCAACCA 480  
Db 421 AACCCAAACACAAACCCAGCAGCCACTACAAAACAAAGCCCAACCAACCAACCAACCA 480  
QY 481 ACCCAATATGATTTTCACTTCCGAGTGTAACTTTGTACCTCGAGCATATGAGCA 540  
Db 481 ACCCAATATGATTTTCACTTCCGAGTGTAACTTTGTACCTCGAGCATATGAGCA 540

QY 541 CAATCCAACTGTCTGGGCTATCTGCAAAAGATACCAAAACAAAAACCAAGAAAGAAAC 600  
Db 541 CAATCCAACTGTCTGGGCTATCTGCAAAAGATACCAAAACAAAAACCAAGAAAGAAAC 600  
QY 601 CACCAACCAAGCCCTTACAAAAAACCACCTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660  
Db 601 CACCAACCAAGCCCTTACAAAAAACCACCTTCAAGACCAACCAAAAAAGATCTCAAACTCA 660  
QY 661 AACCACTAAACCAAGGAAGTACCCACCAACGAGCCCAACCAAGGAGCCCAACCAACAC 720  
Db 661 AACCACTAAACCAAGGAAGTACCCACCAACGAGCCCAACCAAGGAGCCCAACCAACAC 720  
QY 721 CACCAAAACAAACATCACTACACTGCTGCTACCAACCAACCAACCAAGGAAATCCAAAAC 780  
Db 721 CACCAAAACAAACATCACTACACTGCTGCTACCAACCAACCAACCAAGGAAATCCAAAAC 780  
QY 781 CACAAGTCAATTTGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATTTGGAACCTTCCACTCAACCTTCCGAGGCAATCTTAAGCCCTTCTCA 840  
QY 841 AGTCTCCACAACTCGAGCACCACATCAACCCCTCATCTCCACCAACCAACCAACGCGCA 900  
Db 841 AGTCTCCACAACTCGAGCACCACATCAACCCCTCATCTCCACCAACCAACCAACGCGCA 900  
QY 901 GTAGTTATTAAAAA 920  
Db 901 GTAGTTATTAAAAA 920

RESULT 7  
AR148357  
LOCUS AR148357 920 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 7 from patent US 6225091.  
ACCESSION AR148357  
VERSION AR148357.1 GI:15112447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyshyn, M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;  
FEATURES Location/Qualifiers  
source  
1..920  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN

Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACAACTGTCCAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
Db 1 TGCACAACTGTCCAAAACAGGACCAACGCGCTTAAGACACTAGAAAGACCTGGGA 60  
QY 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
Db 61 CACTCTCAATCAATTTATTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGT 120  
QY 121 AGCACAATCAATATTCATATTCGCAATGATAATCTCACTTCACTTATTAATACAGC 180  
Db 121 AGCACAATCAATATTCATATTCGCAATGATAATCTCACTTCACTTATTAATACAGC 180  
QY 181 CATCATATTCATAGCTTCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
Db 181 CATCATATTCATAGCTTCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
QY 241 TGCACAAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 241 TGCACAAAGCCAGATCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
QY 301 ATTCAGCTTCTCAATCTGTCTGAAATTAATCAATCAACCAACCAACCAACCAACCA 360  
Db 301 ATTCAGCTTCTCAATCTGTCTGAAATTAATCAATCAACCAACCAACCAACCAACCA 360  
QY 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420  
Db 361 AACACAGAGGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420  
QY 421 AACCCAAACACAAACCCAGCAGCCACTACAAAACAAAGCCCAACCAACCAACCAACCA 480  
Db 421 AACCCAAACACAAACCCAGCAGCCACTACAAAACAAAGCCCAACCAACCAACCAACCA 480  
QY 481 ACCCAATATGATTTTCACTTCCGAGTGTAACTTTGTACCTCGAGCATATGAGCA 540  
Db 481 ACCCAATATGATTTTCACTTCCGAGTGTAACTTTGTACCTCGAGCATATGAGCA 540

Db 241 TGCACAGCCAGATCAAGACACCAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
QY 301 AATCAGCTTCTCAATCTGTCTGAATTAATCAACAAACCAACCACTATAGCTTCAAC 360  
Db 301 AATCAGCTTCTCAATCTGTCTGAATTAATCAACAAACCAACCACTATAGCTTCAAC 360  
QY 361 AACACAGGAGTCAAGTCAACCTGCAACCCCAACAGTCAAGCTATTAACCAACCAAC 420  
Db 361 AACACAGGAGTCAAGTCAACCTGCAACCCCAACAGTCAAGCTATTAACCAACCAAC 420  
QY 421 AACCCAAACACACCCAGCAAGCCCACTACAAACCAACGCAACCAACCAACCAAC 480  
Db 421 AACCCAAACACACCCAGCAAGCCCACTACAAACCAACGCAACCAACCAACCAAC 480  
QY 481 ACCCAATATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
Db 481 ACCCAATATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
QY 541 CAATCCAACTGCTGGCTATCTGCAAGATATACCAACCAACCAACCAACCAACCAAC 600  
Db 541 CAATCCAACTGCTGGCTATCTGCAAGATATACCAACCAACCAACCAACCAACCAAC 600  
QY 601 CAGCACAGCTTACAAAAAACAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660  
Db 601 CAGCACAGCTTACAAAAAACAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660  
QY 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAACCAACCAACCAACCAAC 720  
Db 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAACCAACCAACCAACCAAC 720  
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
Db 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
QY 781 CACAAGTCAATAGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATAGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840  
QY 841 AGTCTCCAAACATCCGAGACCCCATCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 900  
Db 841 AGTCTCCAAACATCCGAGACCCCATCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 900  
QY 901 GTAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 920  
Db 901 GTAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 920

RESULT 8  
BD081933  
LOCUS  
DEFINITION  
Nucleic acid vaccines encoding G protein of respiratory syncytial virus.  
ACCESSION  
BD081933  
VERSION  
BD081933.1  
KEYWORDS  
JP 2001512662-A/1  
SOURCE  
Respiratory syncytial virus  
ORGANISM  
Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE  
1. (bases 1 to 920)  
Li, X.; Sambhara, S. and Klein, M.H.  
AUTHORS  
Nucleic acid vaccines encoding G protein of respiratory syncytial  
TITLE  
Patent: JP 2001512662-A 1 28-AUG-2001;  
JOURNAL  
CONNAUGHT LABORATORIES LTD  
COMMENT  
OS Respiratory syncytial virus  
PN JP 2001512662-A/1  
PD 28-AUG-2001  
PF 16-JUL-1998 JP 2000503216  
PR 18-JUL-1997 US 08/896442  
PI XIAOMAO LI, SURYAPRAKESH SAMBHARA, MICHEL H KLEIN PC  
C12N15/09, A61K48/00, C07K16/10, C12P21/08, C12Q1/68, G01N33/53, PC  
C12N15/00

CC Nucleic acid vaccines encoding G protein of respiratory CC  
FH Key Location/Qualifiers  
FT CDS (7)..(901).  
FEATURES  
source  
Location/Qualifiers  
1..920  
/organism="Respiratory syncytial virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:12814"  
BASE COUNT 380 a 290 c 95 g 155 t  
ORIGIN  
Query Match 100.0%; Score 920; DB 6; Length 920;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGCACACATGTCCAAAAACAAGGACCAACGCAACCGCTAAGACACACTAGAAAAAGACCTGGCA 60  
Db 1 TGCACACATGTCCAAAAACAAGGACCAACGCAACCGCTAAGACACACTAGAAAAAGACCTGGCA 60  
QY 61 CACTCTCAATCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 120  
Db 61 CACTCTCAATCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 120  
QY 121 AGCACAAATCACATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 180  
Db 121 AGCACAAATCACATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 180  
QY 181 CATCATATTTCATAGCTTCGGCAAAACCAACAAAGTCAACATTAACAACTGCAATCATACAAGA 240  
Db 181 CATCATATTTCATAGCTTCGGCAAAACCAACAAAGTCAACATTAACAACTGCAATCATACAAGA 240  
QY 241 TGCACACGCGAGTCAAGACACAAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
Db 241 TGCACACGCGAGTCAAGACACAAACCCCAACATACCTCACTCAGGATCTCAGCTTGG 300  
QY 301 AATCAGCTTCTCAATCTGTCTGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360  
Db 301 AATCAGCTTCTCAATCTGTCTGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360  
QY 361 AACACAGGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGCTTAAACCAACCAACCAAC 420  
Db 361 AACACAGGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGCTTAAACCAACCAACCAAC 420  
QY 421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACCAACGCAACCAACCAACCAACCAAC 480  
Db 421 AACCCAAACACAAACCCAGCAAGCCCACTACAAACCAACGCAACCAACCAACCAACCAAC 480  
QY 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
Db 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGTACCTGCGAGCATATGCGCAA 540  
QY 541 CAATCCAACTGCTGGCTATCTGCAAGATATACCAACCAACCAACCAACCAACCAACCAAC 600  
Db 541 CAATCCAACTGCTGGCTATCTGCAAGATATACCAACCAACCAACCAACCAACCAACCAAC 600  
QY 601 CAGCACAGCTTACAAAAAACAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660  
Db 601 CAGCACAGCTTACAAAAAACAACCTTCAAGACCAACCAACCAACCAACCAACCTCA 660  
QY 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAACCAACCAACCAACCAACCAAC 720  
Db 661 AACCACTAAACCAAGAGTACCCACCAACCAAGCCCAACCAACCAACCAACCAACCAAC 720  
QY 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
Db 721 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780  
QY 781 CACAAGTCAATAGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840  
Db 781 CACAAGTCAATAGAAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840  
QY 841 AGTCTCCAAACATCCGAGACCCCATCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 900

Db	841	AGTCTCCACAACATCCGAGCACCCATCACACCCCTCATCTCTCCACCCCAACACACGCGCA	900
Qy	901	GTAGTTATTAAAAAAGAAA	920
Db	901	GTAGTTATTAAAAAAGAAA	920
RESULT 9			
LOCUS	HRSRNAG		
DEFINITION	Human respiratory syncytial (RS) virus mRNA for G protein.		
ACCESSION	X17085		
VERSION	X17085.1	GI:60306	
KEYWORDS	g protein.		
SOURCE	Human respiratory syncytial virus		
ORGANISM	Human respiratory syncytial virus		
REFERENCE	Viruses; ssRNA negative-strand viruses; Mononegavirales;		
AUTHORS	Paramyxoviridae; Pneumovirinae; Pneumovirus.		
	1 (bases 1 to 923)		
	Garcia-Barreno,B., Portela,A., Delgado,T., Lopez,J.A. and		
	Melero,J.A.		
TITLE	Frame shift mutations as a novel mechanism for the generation of		
	neutralization resistant mutants of human respiratory syncytial		
	virus		
JOURNAL	EMBO J. 9 (12), 4181-4187 (1990)		
MEDLINE	91065351		
PUBMED	2249671		
FEATURES			
source	Location/Qualifiers		
	1..923		
	/organism="Human respiratory syncytial virus"		
	/mol_type="mRNA"		
	/db_xref="taxon:11250"		
	16..912		
	/codon_start=1		
	/product="G protein"		
	/protein_id="CAA34937.1"		
	/db_xref="GI:60307"		
	/db_xref="SPTREMBL:Q01929"		
	/translation="MSKNQDQRTAKTLTKTMDLVNLIPTSSGLYKLNLSIAQITLS		
	ILAMISTSLIITAILFIASANKHVLTTALIQDATTSTQTSPTTYLTQDPQLGISFS		
	NLSIISQTTTIIASTTPGVKSNLQPTTVTKNTTQITQTPSKPTTKQRKPNKPN		
	NDFHVFVNVPVCSISNNPTCWAICKRIPNKPKGKKTTPKTKPTFKTKKDLKPQ		
	TTPKKEVPTTKPKTEPTTINTKNTITLLTNNNTGNPKLTSQMETFHSSTSEGNLSP		
	QSQVTSFSEHPSQSPSPNTTRQ"		
BASE COUNT	379 a 290 c 98 g 156 t		
ORIGIN			
Query Match	99.1%;	Score 911.8;	DB 14; Length 923;
Best Local Similarity	99.8%;	Pred. No. 1.9e-190;	
Matches 913;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	1	TGCAACATCTCCAAAACACGAGGACCAACCGCGCTTAAGACACACTAGAAAGACCTGGGA	60
Db	9	TGCAACATCTCCAAAACACGAGGACCAACCGCGCTTAAGACACACTAGAAAGACCTGGGA	68
Qy	61	CACCTCTCAATTCATTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTGT	120
Db	69	CACCTCTCAATTCATTATTATTCATATCATCGGGCTTATATAAGTTAAATCTTAAATCTAT	128
Qy	121	AGCACAAATCACATTATCCATTCCTGGCAATGATAAATCTCAACTTCACTTATAATTACGC	180
Db	129	AGCACAAATCACATTATCCATTCCTGGCAATGATAAATCTCAACTTCACTTATAATTACGC	188
Qy	181	CATCATATTCATAGCCTCGGCAACCAACCAAGTCACACTAACCTGCATCATCAAGA	240
Db	189	CATCATATTCATAGCCTCGGCAACCAACCAAGTCACACTAACCTGCATCATCAAGA	248
Qy	241	TGCAACACGCGCATCAAGAACAACACCCCAACATCTCACTCAGATCTCTCAGCTTGG	300
Db	249	TGCAACACGCGCATCAAGAACAACACCCCAACATCTCACTCAGATCTCTCAGCTTGG	308
Qy	301	AATCAGCTTCTCCAATCTGTCTGAAATTCATACAAACACCAACCATATAGCTTCAAC	360

```

source
1. .922
/organism="Human respiratory syncytial virus"
/mol_type="genomic RNA"
/strain="subgroup A"
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"
/db_xref="taxon:11250"
/cell_line="HEP-2"
16. .912
/gene="G"
16. .912
/partial
/gene="G"
/codon_start=1
/product="glycoprotein"
/protein_id="CA83872.1"
/db_xref="GI:485889"
/db_xref="SPTREMBL:O82068"
/translation="ISKNKQRTAKTLEKTDITLNYLLFISGLYKLNLSIAQITLS
ILMIISLTIIITAIIFIASANKVLTITAIQDATSOIKNTPTVLTQDPOLGISF
NLSEITSTQTTIIASTPGVKNLOPTTKNTTTOPTSPKPTKORONPKPN
NDHFEVFNFPSCISNNPTCWAICRIPNKPGRKTTTKPKFTFKTKDLPQ
TKPKVPTTKPTTEPTINTTKNTIITLLTNNNTGNPKLTSOMETTFHSTSEGNLSP
SQVSTSEHPSPSPPTTRQ"
BASE COUNT 378 a 291 c 97 g 156 t
ORIGIN

Query Match 98.8%; Score 909.2; DB 14; Length 922;
Best Local Similarity 99.6%; Pred. No. 7.1e-190;
Matches 914; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGCACACATCTCAAAAACAAAGACCAACGCGCTAGACACTAGAAAGACCTGGGA 60
DB 9 TGCACACATCTCAAAAACAAAGACCAACGCGCTAGACACTAGAAAGACCTGGGA 68
QY 61 CACTCTCAATTTATTTATTCATATCATCGGCTTATATATAGTTAAATCTTAAATCTGT 120
DB 69 CACTCTCAATTTATTTATTTATTCATATCATCGGCTTATATATAGTTAAATCTTAAATCTAT 128
QY 121 AGCACAATCAGATTATCGATTCCTGGCAATGATAATCTCAACTTCATTTAATTTACGC 180
DB 129 AGCACAATCAGATTATCGATTCCTGGCAATGATAATCTCAACTTCATTTAATTTACGC 188
QY 181 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACCACTGCAATCATACAGA 240
DB 189 CATCATATTCATAGCTCGGCAACCAACCAAGTCACACTAACCACTGCAATCATACAGA 248
QY 241 TGCAACAGCGAGATCAAGAACACCAACCCCAACATACCTCCTCAGGATCCTCAGCTTG 300
DB 249 TGCAACAGCGAGATCAAGAACACCAACCCCAACATACCTCCTCAGGATCCTCAGCTTG 308
QY 301 AATCAGCTTCTCAATCTGCTGCAATTTATCATCAACACCAACCACTAGCTTCAAC 360
DB 309 AATCAGCTTCTCAATCTGCTGCAATTTATCATCAACACCAACCACTAGCTTCAAC 368
QY 361 AACACGAGGATCAAGTCAAACTCGCAACCCCAACAGTCAAGTCAAAACCAACCAAC 420
DB 369 AACACGAGGATCAAGTCAAACTCGCAACCCCAACAGTCAAGTCAAAACCAACCAAC 428
QY 421 AACCCAAAACACACCCAGCAAGCCCACTACAAACCAACGCAAAACCAACCAACCAAC 480
DB 429 AACCCAAAACACACCCAGCAAGCCCACTACAAACCAACGCAAAACCAACCAACCAAC 488
QY 481 ACCCAATATGATTTTCACTTCGAAGTGTATTAATTTGACCTTGCGAGCATATGCGAA 540
DB 489 ACCCAATATGATTTTCACTTCGAAGTGTATTAATTTGACCTTGCGAGCATATGCGAA 548
QY 541 CAATCAACCTGCTGGCTATCTGCAAAAGAAATACCAACCAAAACCAAGGAAGAAAC 600
DB 549 CAATCAACCTGCTGGCTATCTGCAAAAGAAATACCAACCAAAACCAAGGAAGAAAC 608
QY 601 CACCAACAGCTTCAAAAACCAACCTTCAAGACACCAAAAGAGATCTCAAACTCA 660
DB 609 CACCAACAGCTTCAAAAACCAACCTTCAAGACACCAAAAGAGATCTCAAACTCA 668

RESULT 11
RSHGLYG
LOCUS
DEFINITION
HUMAN respiratory syncytial virus (subgroup A) attachment protein
M17212
VERSION
M17212.1 GI:333940
KEYWORDS
attachment glycoprotein; surface glycoprotein.
SOURCE
Human respiratory syncytial virus
ORGANISM
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Paramyxoviridae; Pneumovirinae; Pneumovirus.
REFERENCE
1 (bases 1 to 917)
AUTHORS
Johnson, P.R., Spriggs, M.K., Olmsted, R.A. and Collins, P.L.
TITLE
The G glycoprotein of human respiratory syncytial viruses of
subgroups A and B: extensive sequence divergence between
antigenically related proteins
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)
MEDLINE
87289657
PUBMED
2441388
COMMENT
Original source text: Human respiratory syncytial virus (subgroup A,
strain Long), cDNA to viral RNA, clones pUB63, pUB69, and pA3.
The exact 5' end of Long G mRNA was not determined.
FEATURES
source
1. .917
/organism="Human respiratory syncytial virus"
/mol_type="genomic RNA"
/db_xref="taxon:11250"
16. .912
/feature="attachment glycoprotein (G)"
/codon_start=1
/db_xref="GI:333941"
/translation="MSKNKQRTAKTLEKTDITLNYLLFISGLYKLNLSIAQITLS
ILMIISLTIIITAIIFIASANKVLTITAIQDATSOIKNTPTVLTQDPOLGISF
NLSEITSTQTTIIASTPGVKNLOPTTKNTTTOPTSPKPTKORONPKPN
NDHFEVFNFPSCISNNPTCWAICRIPNKPGRKTTTKPKFTFKTKDHPQ
TKPKVPTTKPTTEPTINTTKNTIITLLTNNNTGNPKLTSOMETTFHSTSEGNLSP
SQVSTSEHPSPSPPTTRQ"
BASE COUNT 375 a 290 c 97 g 154 t 1 others
ORIGIN

Query Match 98.0%; Score 901.6; DB 14; Length 917;
Best Local Similarity 99.6%; Pred. No. 3.3e-188;
Matches 904; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCACACATGTCACAAAACAAAGACCAACGCGCTAGACACTAGAAAGACCTGGGAC 61
DB 10 GCACACATGTCACAAAACAAAGACCAACGCGCTAGACACTAGAAAGACCTGGGAC 69
QY 62 ACTCTCAATCATTTATTTATTCATATCATCGGCTTATATATAGTTAAATCTTAAATCTGTA 121

```



RESULT 13  
AR092548  
LOCUS AR092548 894 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 28 from patent US 5998169.  
ACCESSION AR092548  
VERSION AR092548.1 GI:10019302  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 894)  
AUTHORS Klein, M.H., Du, R.-P. and Ewaszyzyn, M.E.  
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
JOURNAL Patent: US 5998169-A 28 07-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..894  
BASE COUNT 364 a 288 c 93 g 149 t  
ORIGIN  
Query Match 97.2%; Score 894; DB 6; Length 894;  
Best Local Similarity 100.0%; Pred. No. 1.6e-186;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 ATGTCCAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 67  
DB 1 ATGTCCAAAACAAAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGACACTCTC 60  
QY 68 AATCATTTATTATTCATATCATCTGGCTTATATTAAGTAAATCTTAAATCTGTAGCAAA 127  
DB 61 AATCATTTATTATTCATATCATCTGGCTTATATTAAGTAAATCTTAAATCTGTAGCAAA 120  
QY 128 ATCATATTATCCATCTTGGCAATGATATCTCACTTCACTTATATATACAGCCATCATTA 187  
DB 121 ATCATATTATCCATCTTGGCAATGATATCTCACTTCACTTATATATACAGCCATCATTA 180  
QY 188 TTCTAGCTCGGCAAAACCAAGTCACTAACAATCAATCAATCAATCAATCAATCAATCAAT 247  
DB 181 TTCTAGCTCGGCAAAACCAAGTCACTAACAATCAATCAATCAATCAATCAATCAATCAAT 240  
QY 248 AGCCAGATCAAGACACAAACCCCACTACTCTCAGGATCTCTAGTCTGAGTCAATCAGC 307  
DB 241 AGCCAGATCAAGACACAAACCCCACTACTCTCAGGATCTCTAGTCTGAGTCAATCAGC 300  
QY 308 TTCTCCAACTCTGTGAATTTATCATCAAAACCAACCACTACTAGCTTCAACCAACCA 367  
DB 301 TTCTCCAACTCTGTGAATTTATCATCAAAACCAACCACTACTAGCTTCAACCAACCA 360  
QY 368 GGAGTCAAGTCAAACTCTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 427  
DB 361 GGAGTCAAGTCAAACTCTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420  
QY 428 ACACACCCAGAGCCACTTACAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAAT 487  
DB 421 ACACACCCAGAGCCACTTACAAACACGCGCAACCAACCAACCAACCAACCAACCAACCAAT 480  
QY 488 AATGATTTTCACTTCCGAAGTGTGTTAACTTTGTACCTCTGAGCATATGCGACCAATCCA 547  
DB 481 AATGATTTTCACTTCCGAAGTGTGTTAACTTTGTACCTCTGAGCATATGCGACCAATCCA 540  
QY 548 ACTGTCTGGGTATCTCAAAAAGATACCAACCAAAAACCGAGAAAGAAACCAACCAACC 607  
DB 541 ACTGTCTGGGTATCTCAAAAAGATACCAACCAAAAACCGAGAAAGAAACCAACCAACC 600  
QY 608 AGCCTTCAAAAACCAACCTTCAAGACACCAACCAAAAAGATCTCAAACTTCAAACT 667  
DB 601 AGCCTTCAAAAACCAACCTTCAAGACACCAACCAAAAAGATCTCAAACTTCAAACT 660  
QY 668 AAACCAAGGAAGTACCCACCAACCAAGCCCAAGAGAGGCAACCATCAACACCAACCA 727

Db 661 AAACCAAGGAAGTACCCACCAAGCCCAAGAGCCCAAGAGACCAACCATCAACACCAAA 720  
QY 728 ACAAACATCAAACTACTGTCTCAACCAACCAACCAAGAAATCCAAATCTCAAACT 787  
Db 721 ACAAACATCAAACTACTGTCTCAACCAACCAACCAAGAAATCCAAATCTCAAACT 780  
QY 788 CAATGGAAACCTTCCACTCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 847  
Db 781 CAATGGAAACCTTCCACTCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCC 840  
QY 848 ACAAACATCGAGCACCCATCAACAAACCTCTATCTCCACCAACCAACCAACGCGCAG 901  
Db 841 ACAAACATCGAGCACCCATCAACAAACCTCTATCTCCACCAACCAACCAACGCGCAG 894  
RESULT 14  
RSHICE 8510 bp ss-RNA linear VRL 29-NOV-2000  
LOCUS Human respiratory syncytial virus nonstructural protein (1C),  
DEFINITION nonstructural protein (1b), major nucleocapsid (N), phosphoprotein  
(P), protein (M), 1A (1A), G (G), protein (F) and  
envelope-associated protein (22K) gene, complete cds.  
ACCESSION M11486 K01459 K03348 K03349 M11217 M11244 M11487 M11505  
VERSION M11514 M11631 M12966  
KEYWORDS M11486.1 GI:333925  
envelope-associated protein; fusion glycoprotein; major  
nucleocapsid protein; major surface glycoprotein; matrix protein;  
nonstructural protein; phosphoprotein.  
SOURCE Human respiratory syncytial virus  
ORGANISM Human respiratory syncytial virus  
REFERENCE 1 (bases 1085 to 2220; 5254 to 5544)  
AUTHORS Paranyxoviridae; Pneumovirinae; Pneumovirus.  
TITLE Elango, N. and Venkatesan, S.  
JOURNAL Amino acid sequence of human respiratory syncytial virus  
MEDLINE Nucleic Acids Res. 11 (17), 5941-5951 (1983)  
PUBMED 83299261  
REFERENCE 2 (bases 3211 to 4157)  
AUTHORS Satake, M. and Venkatesan, S.  
TITLE Nucleotide sequence of the gene encoding respiratory syncytial  
virus matrix protein  
JOURNAL J. Virol. 50 (1), 92-99 (1984)  
MEDLINE 84138836  
PUBMED 6699948  
REFERENCE 3 (bases 2288 to 3191)  
AUTHORS Satake, M., Elango, N. and Venkatesan, S.  
TITLE Sequence analysis of the respiratory syncytial virus phosphoprotein  
JOURNAL J. Virol. 52 (3), 991-994 (1984)  
MEDLINE 85033973  
PUBMED 6548527  
REFERENCE 4 (bases 5602 to 7500)  
AUTHORS Collins, P.L., Huang, Y.T. and Wertz, G.W.  
TITLE Nucleotide sequence of the gene encoding the fusion (F)  
glycoprotein of human respiratory syncytial virus  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)  
MEDLINE 85088471  
PUBMED 6096849  
REFERENCE 5 (bases 7551 to 8510)  
AUTHORS Collins, P.L. and Wertz, G.W.  
TITLE The envelope-associated 22K protein of human respiratory syncytial  
virus: nucleotide sequence of the mRNA and a related polytranscript  
JOURNAL J. Virol. 54 (1), 65-71 (1985)  
MEDLINE 85135082  
PUBMED 3838351  
REFERENCE 6 (bases 8 to 1050; 7554 to 8506)  
AUTHORS Elango, N., Satake, M. and Venkatesan, S.  
TITLE mRNA sequence of three respiratory syncytial virus genes encoding  
two nonstructural proteins and a 22K structural protein  
JOURNAL J. Virol. 55 (1), 101-110 (1985)  
MEDLINE 85237684



PUBMED REFERENCE AUTHORS	4009789 7 (bases 5602 to 7423) Elango,N., Satake,M., Coligan,J.E., Norrby,E., Camargo,E. and Venkatesan,S.	gene	/db_xref="taxon:11250" /lab_host="HEP-2 cell" 1..528 /gene="1C" 1..528 /gene="1C" 55..474 /gene="1C" /codon_start=1 /product="nonstructural protein" /protein_id="AAB59850.1" /db_xref="GI:333926"
TITLE	Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit	mRNA	/translation="MGSNSLSMIKVRQLQMLFNDVALLKTCYTDKLIHLTNALAKA VLIHTKLNGIVFVHVTISDPCPNNIYVKSNTFMPVLQNGGYIWEWMLTHCSQPN GLLDNDCEIKFSKSLSDSTWNTNMQLSLLGFDLNP" 171
JOURNAL MEDLINE PUBMED	85215565 2987829 8 (bases 4630 to 5543) Satake,M., Coligan,J.E., Elango,N., Norrby,E. and Venkatesan,S. Respiratory syncytial virus envelope glycoprotein (G) has a novel structure	variation	/gene="1C" /note="c in [11]; t in [6]" /replace="t" 552..1050 /gene="1B" 552..1050 /gene="1B" 584..958 /gene="1B" /codon_start=1 /product="nonstructural protein" /protein_id="AAB59851.1" /db_xref="GI:333927"
REFERENCE AUTHORS	86067198 4069997 9 (bases 4627 to 5544) Wertz,G.W., Collins,P.L., Huang,Y., Gruber,C., Levine,S. and Ball,L.A.	gene	/translation="MDTTHNDNTPORLMTDMRPLSLRTITSLTRDITTHKFIYLIN HECIVRLDEKQATFETLVNYEMKLLHKVGSYKKYKTYETNTKYTFPMPPIFNDGPF LECGIKPTKPTPIIYKYDLNP" 747
TITLE	Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein	mRNA	/gene="1B" /note="a in [11]; g in [6]" /replace="g" 769
JOURNAL MEDLINE PUBMED	85216636 3858865 10 (bases 4173 to 4571) Collins,P.L. and Wertz,G.W. The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript	variation	/gene="1B" /note="g in [11]; a in [6]" /replace="a" 1081..2277 /gene="N" 1081..2277 /gene="N" 1096..2271 /gene="N" /codon_start=1 /product="major nucleocapsid" /protein_id="AAB59852.1" /db_xref="GI:333928"
REFERENCE AUTHORS	86098645 3879976 11 (bases 1 to 528; 552 to 1050) Collins,P.L. and Wertz,G.W. Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus	gene	/translation="MALSQVKNLNTLNKQDLLSSSKYTIQRSTGDSIDTPNYDVQKH NKLQCMGLITEDANHKFTGLIGMLYMRGLGREDTIKLRDAGYHVKANGVDVTHRQ DINGKMEFVLTSLTETIQINIEISRSKYKMKLMEGVAPEYRHDSDDCCMI LCIAALVTTLAAGDSRGLTAVIRRNANLVKNEMKRYKGLLPKDIANSFYEFKHPH FIDVPHFGIAQSGTSRGSRRVEGIFAGLPMNAYGAGQVMLNGLAKSVKIMLGHAS VQAMEQVVEYVEYAQKLGAGAGTHILNPKASLLSLTQPFHFSVVLGNAAGLIGIM GEYRGTPRNQDLIDAAKAYASQLKENGVINYSVLDTABELEATKHQLNPKNDQVEL" 2289..3196 /gene="p" 2289..3196 /gene="p" 2302..3027 /gene="p" /codon_start=1 /product="phosphoprotein" /protein_id="AAB59853.1" /db_xref="GI:333929"
JOURNAL MEDLINE PUBMED	86045905 2998021 12 (bases 1081 to 2277) Collins,P.L., Anderson,K., Langer,S.J. and Wertz,G.W. Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus	mRNA	/translation="MEKFAPEFHGEDANNRATKFLSIKGFSTPKDFKPKKDSIISVN SIDIEVTKESPTISNTIINFNETDGTAGNPKYQKPLVSFKEDPTPSNPSKLY KETIETFDNNEESSYSYERINDQTDNDNIARLDIKLEIKLGLMLHTLVASAGPT SARDGIRDAMIGLREEMIEKIRTEALMTNDRLEAMARLNEESEKWKADTSDVSLNP 92.8%; Score 853.4; DB 14; Length 8510;
TITLE	Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	CDS	
JOURNAL MEDLINE PUBMED	85301974 3839952 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578) Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camargo,E. and Coelingh,K.V. Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	variation	
REFERENCE AUTHORS	86259643 3460060 14 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578) Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camargo,E. and Coelingh,K.V. Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	gene	
TITLE	Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	mRNA	
JOURNAL MEDLINE PUBMED	86259643 3460060 14 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578) Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camargo,E. and Coelingh,K.V. Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order	CDS	
COMMENT	Reprint for [2] and clean copy sequence for [12], [5] kindly provided by P.Collins, 21-FEB-1986. Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown.		
FEATURES source	Location/Qualifiers 1..8510 /organism="Human respiratory syncytial virus" /mol_type="genomic RNA" /strain="A2"	Query Match	





Qy	781	CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA	840
Db	5461	CACAAGTCAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA	5520
Qy	841	AGTCTCCACACATCCGAGCACCCTCATCAACCTCTCTCCACCCCAACACACGCCA	900
Db	5521	AGTCTCTACACATCCGAGTACCCTCATCAACCTCTCTCCACCCCAACACACGCCA	5580
Qy	901	GTAGTTATTAAAAAAA	919
Db	5581	GTAGTTACTTAAAAACATA	5599

Search completed: October 30, 2003, 00:01:07  
Job time : 3766.15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:32:25 ; Search time 70.283 Seconds  
(without alignments)  
1094.143 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKQRTAKTLEKTWDTL.....VSTTSEHPSPSSPPNTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537	99.5	298	12 Q01929	Q01929 human respi
2	1533	99.3	298	12 Q82068	Q82068 human respi
3	1393	90.2	298	12 Q09634	Q09634 human respi
4	1389	90.0	298	12 Q09719	Q09719 respiratory
5	1365	88.4	297	12 Q9YVB5	Q9YVB5 human respi
6	1364	88.3	297	12 Q82066	Q82066 human respi
7	1356	87.8	298	12 Q9YVB4	Q9YVB4 human respi
8	1354	87.7	297	12 Q82067	Q82067 human respi
9	1345	87.1	297	12 Q82071	Q82071 human respi
10	1344	87.0	297	12 Q82058	Q82058 human respi
11	1339	86.7	297	12 Q91947	Q91947 human respi
12	1336	86.5	293	12 Q9YVB3	Q9YVB3 human respi
13	1334	86.4	295	12 Q86359	Q86359 respiratory
14	1334	86.4	295	12 Q86360	Q86360 respiratory
15	1334	86.4	298	12 Q82065	Q82065 human respi
16	1331	86.2	297	12 Q82074	Q82074 human respi

17	1331	86.2	297	12 Q82064	Q82064 human respi
18	1328	86.0	297	12 Q82063	Q82063 human respi
19	1326	85.9	292	12 Q9YVB2	Q9YVB2 human respi
20	1326	85.9	298	12 Q82077	Q82077 human respi
21	1325	85.8	298	12 Q82060	Q82060 human respi
22	1324	85.8	297	12 Q91946	Q91946 human respi
23	1320	85.5	297	12 Q82070	Q82070 human respi
24	1319	85.4	293	12 Q9YVB1	Q9YVB1 human respi
25	1319	85.4	295	12 Q86361	Q86361 respiratory
26	1319	85.4	297	12 Q82079	Q82079 human respi
27	1319	85.4	297	12 Q82057	Q82057 human respi
28	1318	85.4	297	12 Q9YVC8	Q9YVC8 human respi
29	1317	85.3	298	12 Q82075	Q82075 human respi
30	1317	85.3	298	12 Q82062	Q82062 human respi
31	1317	85.3	298	12 Q82056	Q82056 human respi
32	1316	85.2	297	12 Q82072	Q82072 human respi
33	1315	85.2	295	12 Q86357	Q86357 respiratory
34	1314	85.1	297	12 Q91944	Q91944 human respi
35	1314	85.1	297	12 Q82078	Q82078 human respi
36	1314	85.1	298	12 Q82073	Q82073 human respi
37	1314	85.1	298	12 Q82076	Q82076 human respi
38	1312	85.0	298	12 O11978	O11978 human respi
39	1311	84.9	292	12 Q9YVB0	Q9YVB0 human respi
40	1311	84.9	298	12 Q91945	Q91945 human respi
41	1308	84.7	296	12 Q86358	Q86358 respiratory
42	1306	84.6	298	12 Q8JQU4	Q8JQU4 human respi
43	1306	84.6	298	12 Q82069	Q82069 human respi
44	1305	84.5	298	12 Q82061	Q82061 human respi
45	1298	84.1	295	12 Q86356	Q86356 respiratory

## ALIGNMENTS

### RESULT 1

Q01929 PRELIMINARY; PRT; 298 AA.  
 ID Q01929  
 AC Q01929;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE G protein.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91065351; PubMed=2249671;  
 RA Garcia-Barrero B., Portela A., Delgado T., Lopez J.A., Melero J.A.;  
 RT "Frame shift mutations as a novel mechanism for the generation of  
 RT newralization resistant mutants of human respiratory syncytial  
 RT virus.";  
 RL EMBO J. 9:4181-4187(1990).  
 DR EMBL; X17085; CAA34937.1; -;  
 DR InterPro; IPR000925; Glycoprot G  
 DR InterPro; IPR006162; Ppantne attach.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
 SQ SEQUENCE 298 AA; F8CD4213D97C2952 CRC64;

Query Match 99.5%; Score 1537; DB 12; Length 298;  
 Best Local Similarity 99.3%; Pred. No. 1e-105;  
 Matches 296; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMITSLITAI 60  
 Db 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMITSLITAI 60  
 Qy 61 FIASANHKVLTATTAIQDATSQIKNTTPTVLTQDPQLGIFSFLSITSTTTILASTTP 120  
 Db 61 FIASANHKVLTATTAIQDATSQIKNTTPTVLTQDPQLGIFSFLSITSTTTILASTTP 120

```

QY 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
Db 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
QY 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 2
Q82068 PRELIMINARY; PRT; 298 AA.
AC Q82068;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33429; CAA83872.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne_attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
FT NON TER 1
SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 99.3%; Score 1533; DB 12; Length 298;
Best Local Similarity 99.0%; Pred. No. 2e-105;
Matches 295; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
Db 1 ISKNKQDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
QY 61 FIASANKVLTITAIQDATSQIKNTPTLYLTQDPQLGIFSFLSLSITSTQTTILASTTP 120
Db 61 FIASANKVLTITAIQDATSQIKNTPTLYLTQDPQLGIFSFLSLSITSTQTTILASTTP 120
QY 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
Db 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
QY 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 3
Q09634 PRELIMINARY; PRT; 298 AA.
ID Q09634

```

```

AC Q09634;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Attachment protein (G).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Flows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2;
RA Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39662; AAC57026.1; -
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne_attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match 90.2%; Score 1393; DB 12; Length 298;
Best Local Similarity 91.3%; Pred. No. 4.3e-95;
Matches 272; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKNKQDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
Db 1 MSKNKQDQRTAKTLEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMIISTLIITAI 60
QY 61 FIASANKVLTITAIQDATSQIKNTPTLYLTQDPQLGIFSFLSLSITSTQTTILASTTP 120
Db 61 FIASANKVLTITAIQDATSQIKNTPTLYLTQDPQLGIFSFLSLSITSTQTTILASTTP 120
QY 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
Db 121 GVSNNLOPTTVTKNTTTTQPSKPTTKQKQKPNKPNNDPHEVEFVPCISCSNNP 180
QY 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
Db 181 TCWAIKCRIPNKKPKGKTTTKTKPTTKTKKDLKPQTTKPKKEVPTTKPTPEPTINTTK 240
QY 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298
Db 241 TMTTLLTNNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 4
Q09719 PRELIMINARY; PRT; 298 AA.
ID Q09719
AC Q09719;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Attachment glycoprotein (G).
GN G.
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1C;

```

```

RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL vaccine 14:1637-1646 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S2 ts1c;
RC Easton A.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39661; AAC57036.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32775 MW; 67FA043682FA450 CRC64;

Query Match 90.0%; Score 1389; DB 12; Length 298;
Best Local Similarity 90.9%; Pred. No. 8.4e-95;
Matches 271; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
DB 121 SVKSTLQSTTVTKNTTTTKIQPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
QY 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
QY 241 TNIITTLTNTTGNPKLTSQMETFHTSTSGNLSPSQVSTTSEHPSPSPNTTQ 298
DB 241 TNIITTLTNTTGNPKLTSQMETFHTSTSGNLSPSQVSTTSEHPSPSPNTTQ 298

RESULT 5
QYQVVB5 PRELIMINARY; PRT; 297 AA.
AC QYVVB5
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VW2780;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Murson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932 (1998).
DR EMBL; AF065405; AAO2941.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 297 AA; 32572 MW; 5C10E719A05EF5C1 CRC64;

Query Match 88.4%; Score 1365; DB 12; Length 297;

```

```

Best Local Similarity 90.2%; Pred. No. 4.9e-93;
Matches 268; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
DB 121 GVKSTLQSTTVTKNTTTTQAPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
QY 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
QY 241 TNIITTLTNTTGNPKLTSQMETFHTSTSGNLSPSQVSTTSEHPSPSPNTTQ 297
DB 241 TNIITTLTNTTGNPKLTSQMETFHTSTSGNLSPSQVSTTSEHPSPSPNTTQ 297

RESULT 6
QYQVVB5 PRELIMINARY; PRT; 297 AA.
AC QYVVB5
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabraile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459 (1994).
DR EMBL; Z33427; CAA83870.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 297 AA; 32589 MW; F210F8C68193F5DD CRC64;

Query Match 88.3%; Score 1364; DB 12; Length 297;
Best Local Similarity 89.9%; Pred. No. 5.8e-93;
Matches 267; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
QY 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
DB 61 FIASANKHVLTTLAIQDATSQIKNTTPTLTQDPQLGISPSNLSITTSQTITLSTTP 120
QY 121 GVKSNLQPTTVKTKNTTTTQTPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
DB 121 GVKSTLQSTTVTKNTTTTQAPSKPTTKQKQPNKPNNDFFHFEVFPVCSICSNP 180
QY 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240
DB 181 TCWAIKRIKPNKPGKKTTPKTKPTTKKOLKPKQTTKPKVPTTKPTPEPTINTTK 240

```

```

Db 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKKSEVPTTKPTBPTINTTK 240
QY 241 TMTITLLTNNTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 297
Db 241 TMTITLLTNNTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTSR 297

RESULT 7
ID Q9YVB4 PRELIMINARY; PRT; 298 AA.
AC Q9YVB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV5222;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065406; AA02942.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
SQ SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 87.8%; Score 1356; DB 12; Length 298;
Best Local Similarity 88.9%; Pred. No. 2.3e-92;
Matches 265; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60
Db 1 MSKNKDQRTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60

QY 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIFSNLSSETTSQTITLASTTP 120
Db 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIFSNLSSETTSQTITLASTTP 120

QY 121 GVKSNLOPTTVKTKNTTTTQPSKPTTKORONKPNKPNDFHFEVNFVPCSISSNP 180
Db 121 SVESTLQSTTVKTKNTTTTQPSKPTTKORONKPNKPNDFHFEVNFVPCSISSNP 180

QY 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKKSEVPTTKPTBPTINTTK 240
Db 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKKSEVPTTKPTBPTINTTK 240

QY 241 TMTITLLTNNTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 298
Db 241 PNIRTTLLTNNTGSPHTSKETLHSTSEGNLSPSQVSTTSEHPSQSPSPNTNQ 298

RESULT 8
ID Q82071 PRELIMINARY; PRT; 297 AA.
AC Q82071;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94335057; PubMed=8057427;
RX Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33428; CAA83871.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
SQ SEQUENCE 297 AA; 32555 MW; 9D0A69BECAE8B6C8 CRC64;

Query Match 87.7%; Score 1354; DB 12; Length 297;
Best Local Similarity 89.2%; Pred. No. 3.2e-92;
Matches 265; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60
Db 1 MSKNKDQRTAKTLEKWTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLITAI 60

QY 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIFSNLSSETTSQTITLASTTP 120
Db 61 FIASANKVTLTALIODATSOIKNTPTTYPOLGIFSNLSSETTSQTITLASTTP 120

QY 121 GVKSNLOPTTVKTKNTTTTQPSKPTTKORONKPNKPNDFHFEVNFVPCSISSNP 180
Db 121 GVKSLQSTTVKTKNTTTTQPSKPTTKORONKPNKPNDFHFEVNFVPCSISSNP 180

QY 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKKSEVPTTKPTBPTINTTK 240
Db 181 TCWAICKRIPNKKPKGKRTTKTKPLTKTKDQPKQTKKSEVPTTKPTBPTINTTK 240

QY 241 TMTITLLTNNTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 297
Db 241 TMTITLLTNNTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTSR 297

RESULT 9
ID Q82067 PRELIMINARY; PRT; 297 AA.
AC Q82067;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94335057; PubMed=8057427;
RX Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33428; CAA83871.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
SQ SEQUENCE 297 AA; 32555 MW; 9D0A69BECAE8B6C8 CRC64;

```

Query Match 87.1%; Score 1345; DB 12; Length 297;  
Best Local Similarity 88.9%; Pred. No. 1.5e-91;  
Matches 264; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIAI 60

QY 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSOPTTILASTTP 120  
DB 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSLITILDSTTP 120

QY 121 GVKSNLQPTTKVTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 GVKSTLQSTTVGKNTTTTQDQPKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297  
DB 241 TNITITLLTNNTGNPKLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297

RESULT 10  
Q82058 PRELIMINARY; PRT; 297 AA.

ID Q82058  
AC Q82058  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE (MAD-1-89) subgroup A, G glycoprotein.  
GN G.

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=subgroup A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabraile S., Russi J.,  
RA Hortal A., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Melero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL: Z33456; CAA83879.1; --  
DR InterPro; IPR000925; Glycoprot.G.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHENE; 1.  
SQ SEQUENCE 297 AA; 32741 MW; 0B567A174BF64964 CRC64;

Query Match 87.0%; Score 1344; DB 12; Length 297;  
Best Local Similarity 88.5%; Pred. No. 1.7e-91;  
Matches 262; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIAI 60

QY 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSOPTTILASTTP 120  
DB 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSOPTTILASTTP 120

QY 121 GVKSNLQPTTKVTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 SAESTPQSTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 296  
DB 241 TNITITLLTNNTGNPKLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 11  
Q91947 PRELIMINARY; PRT; 297 AA.

ID Q91947  
AC Q91947  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Attachment protein.  
GN G.

OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CH34;  
RX MEDLINE=38418504; PubMed=9747732;  
RA Peret T.C.T., Hall C.B., Schnabel K.C., Golub J.A., Anderson L.J.;  
RT "Circulation Patterns of Genetically Distinct Group A and B Strains of  
RT Human Respiratory Syncytial Virus in a Community.";  
RL J. Gen. Virol. 79:2221-2229(1998).  
DR EMBL: AF065257; AAC36327.1; --  
DR InterPro; IPR000925; Glycoprot.G.  
DR InterPro; IPR006162; Ppantne attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHENE; 1.  
SQ SEQUENCE 297 AA; 32564 MW; 377C7CBCDD0A3645 CRC64;

Query Match 86.7%; Score 1339; DB 12; Length 297;  
Best Local Similarity 88.2%; Pred. No. 4.1e-91;  
Matches 262; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKSAVAQITLSILAMIISTSLIIAI 60

QY 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSOPTTILASTTP 120  
DB 61 FIASANHKVLTITAIQDQTSQIKNTTPTLTQDQGLGSPNSLSEITSLITILDSTTP 120

QY 121 GVKSNLQPTTKVTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180  
DB 121 GVKSTLQSTTVGKNTTTTQDQPKPTTKQKQKPNKPNNDFFHFEVFNVPSCSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKPTKTKKOLKPKQPTTKSKEVPTTKTEPTINTTK 240

QY 241 TNITITLLTNNTGNPKLTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297  
DB 241 TDITITPTTNTNPNELTTSQMETPHSTSSGNLSPSOVSTTSEHPSQSPSPNTT 297

RESULT 12  
Q9YVB3 PRELIMINARY; PRT; 293 AA.

ID Q9YVB3  
AC Q9YVB3  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Attachment glycoprotein G (fragment).  
GN Human respiratory syncytial virus.  
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

```

OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV6973;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065407; AAD02943.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
FT NON TER 1
SQ SEQUENCE 293 AA; 32125 MW; 17B5B43396A63CCF CRC64;

Query Match 86.5%; Score 1336; DB 12; Length 293;
Best Local Similarity 90.1%; Pred. No. 6.6e-91;
Matches 264; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

QY 6 DQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPIASA 65
DB 1 DQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPIASA 60

QY 66 NHKVTLTALIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGVKSN 125
DB 61 NHRVSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGVKST 120

QY 126 LQPTTKTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCISCSNNPTCAI 195
DB 121 LQSTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCISCSNNPTCAI 180

QY 186 CKRIPNKKPKKTKTKPTKTKPTKTKDLPQTKPKKPTTKPKKEVPTTKPEPTINTKNTIT 245
DB 181 CKRIPNKKPKKTKTKPTKTKPTKTKDLPQTKPKKPTTKPKKEVPTTKPEPTINTKNTIT 240

QY 246 TLTNTNTGNPKLTSMQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTQ 298
DB 241 TLTNTNTRNPELTSMQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTQ 293

RESULT 13
Q86359 PRELIMINARY; PRT; 295 AA.
AC Q86359;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73353; CAA51764.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
FT NON TER 1
SQ SEQUENCE 293 AA; 32125 MW; 17B5B43396A63CCF CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.5%; Pred. No. 9.4e-91;
Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 122
DB 61 ASANHKITSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 120

```

```

FT NON TER 1
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.4%; Pred. No. 9.4e-91;
Matches 260; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KTKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 122
DB 61 ASANHKVTLTALIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 120

QY 123 KSNLQPTTKTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCISCSNNPTC 182
DB 121 ESTPQSTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCISCSNNPTC 180

QY 183 WAICKRIPNKKPKKTKTKPTKTKPTKTKDLPQTKPKKPTTKPKKEVPTTKPEPTINTKTN 242
DB 181 WAICKRIPNKKPKKTKTKPTKTKPTKTKDLPQTKPKKPTTKPKKEVPTTKPEPTINTKTN 240

QY 243 ITTTLTNTNTPKLTSMQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 296
DB 241 ITTTLTNTNTPKLTSMQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 294

RESULT 14
Q86360 PRELIMINARY; PRT; 295 AA.
AC Q86360;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73354; CAA51765.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHINE; 1.
FT NON TER 1
SQ SEQUENCE 295 AA; 32306 MW; 671DCF044157D284 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
Best Local Similarity 88.5%; Pred. No. 9.4e-91;
Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 62
DB 1 KNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTLSIIITAIPI 60

QY 63 ASANHKVTLTALIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 122
DB 61 ASANHKITSTTTIIQDATSQIKNTTPTLTQDQGLGIFSNLSSEITSSQTTILASTTPGV 120

```



```

QY 123 KSNLQPTTVTKNTTQTOTQSPKPTTKQKQKPPKPNNDFFHEVFNFPVPCSCSNPTC 182
DB 121 KSTLQSTTVGTNTTTOAQPNKPTTKQKQKPPKPNNDFFHEVFNFPVPCSCSNPTC 180
QY 183 WAICKRIPNKKPGKTKTKPTTKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTKN 242
DB 181 WAICKRIPNKKPGKTKTKPTTKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTKN 240
QY 243 ITTTLTNNNTGNPKLTQMETPHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 297
DB 241 ITTTLTNNTRNPELTQMETPHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 295

RESULT 15
Q82065
ID Q82065 PRELIMINARY; PRT; 298 AA.
AC Q82065;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5455(1994).
DR EMBL; Z33426; CAA83869.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1
DR PROSITE; PS00012; PROSHPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32729 MW; 4E890FFD43B45744 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 298;
Best Local Similarity 88.2%; Pred. No. 9.5e-91;
Matches 261; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTDNLNHLFISSGLYKLNKLSVAQITLSLAMIISTSLIITAI 60
DB 1 MSKTKDQRTAKTLEKTDNLNHLFISSGLYKLNKLSVAQITLSLAMIISTSLIITAI 60
QY 61 FIASANKHVKLTTLTIQDATSQIKNTPTLTODPOLGISFNSLSEITSOITILASTTP 120
DB 61 FIASANKHVKLTTLTIQDATSQIKNTPTLTQNPQLGISFNSLSEITSOITILASTTP 120
QY 121 GVKSNLQPTTVTKNTTTTQPSKPTTKQKQKPPKPNNDFFHEVFNFPVPCSCSNNP 180
DB 121 SAESTFLSTTVTKNTTTTQPSKPTTKQKQKPPKPNNDFFHEVFNFPVPCSCSNNP 180
QY 181 TCWAICKRIPNKKPGKTKTKPTTKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTK 240
DB 181 TCWAICKRIPNKKPGKTKTKPTTKPTTKKOLKPKQTKPKKVPPTTKPTTEPTINTTK 240
QY 241 TNIITTLTNNNTGNPKLTQMETPHSTSSGNLSPSQVSTTSBHPSPSPSPPTTR 296
DB 241 TNIITTLTNNTPHTSOKETLHSTSDGNLSPSQVSTTSBHPSPSPSPPTTR 296

```

Search completed: October 29, 2003, 17:40:52  
Job time : 75.283 secs

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	1530	99.1	298	1	VGLG_HRSVL	P20895 human respi
2	1418	91.8	298	1	VGLG_HRSVA	P03422 human respi
3	1343	87.0	297	1	VGLG_HRSV3	P27022 human respi
4	1343	87.0	297	1	VGLG_HRSV6	P27025 human respi
5	1328	86.0	297	1	VGLG_HRSV7	P27026 human respi
6	1324	85.8	297	1	VGLG_HRSV4	P27022 human respi
7	1313	85.0	298	1	VGLG_HRSV5	P27024 human respi
8	1307	84.7	297	1	VGLG_HRSV2	P27021 human respi
9	767.5	49.7	292	1	VGLG_HRSV8	P23041 human respi
10	761	49.3	292	1	VGLG_HRSV1	P20896 human respi
11	309.5	20.0	263	1	VGLG_BRSV1	O10683 bovine resp
12	308.5	20.0	263	1	VGLG_BRSV4	O10684 bovine resp
13	299	19.4	257	1	VGLG_BRSV	P22261 bovine resp
14	294.5	19.1	263	1	VGLG_BRSV2	O10685 bovine resp
15	295.5	19.1	263	1	VGLG_BRSVW	O10687 bovine resp
16	285	18.5	257	1	VGLG_BRSVS	O10687 bovine resp
17	284	18.4	263	1	VGLG_BRSV	O10686 bovine resp
18	279	18.1	257	1	VGLG_BRSVW	Q86695 ovine respi
19	270	17.5	257	1	VGLG_BRSVL	Q65706 bovine resp
20	267	17.3	257	1	VGLG_BRSVR	Q09495 bovine resp
21	208	13.5	307	1	SGS3_DROME	Q84183 bovine resp
22	198.5	12.9	2812	1	ZAN_HUMAN	P02840 drosophila
23	194	12.6	662	1	MUC1_XENLA	Q9Y493 homo sapien
24	193	12.5	5179	1	MUC2_HUMAN	Q05049 xenopus lae
25	191	12.4	1161	1	DANA_YEAST	Q02817 homo sapien
26	183	11.9	217	1	SGS3_DROSI	P47179 saccharomyc
27	171.5	11.1	2282	1	ZAN_RABIT	P13729 drosophila
28	170	11.0	5703	1	MUSB_HUMAN	P57999 oryctolagus
29	169.5	11.0	1367	1	AMTH_YEAST	Q9Hc84 homo sapien
30	164.5	10.7	2476	1	ZAN_PIG	P08640 saccharomyc
31	160	10.4	263	1	SGS3_DROYA	Q28983 sus scrofa
32	159.5	10.3	907	1	VGLP_BEV	P13728 drosophila
33	157.5	10.2	5376	1	ZAN_MOUSE	P03200 epstein-bar
						Q88799 mus musculus



```
CC CC HEMAGGLUTININATING ACTIVITIES.
CC CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; JQ1205; JQ1205.
DR DR InterPro; IPR000925; Glycoprot G.
DR DR Pfam; PF00802; Glycoprotein G; 1.
DR DR Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32525 MW; 48448F9E091E1802 CRC64;

Query Match 87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.6%; Pred. No. 1.2e-80;
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSXNKDQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSXNKDQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIIVAI 60
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITTSQTTILASTTP 120
DB 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITTSQTTILASTTP 120
QY 121 GVKSNLQPTTKTKNTTTTQPSKPTTKQKQKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 GVKSNTLQSTTVGKNTTTTQAPNKPPTTKQKQKPNKPNNDHFEVFNVPSCISNNP 180
QY 181 TCWAICKRIENKPKGKTTTKPKPTKTKKDLKPQTTKPKKEVPTTKPTBPTINTTK 240
DB 181 TCWAICKRIENKPKGKTTTKPKPTKTKKDLKPQTTKPKKEVPTTKPTBPTINTTK 240
QY 241 TNIITLLTNTTGNPKLTSQMETFSTSGEGLSPSQVSTTSEHPSQSPSPNTTR 297
DB 241 TNIITLLTNTTGNPKLTSQMETFSTSGEGLSPSQVSTTSEHPSQSPSPNTTR 297

RESULT 4
VGLG_HRSV6 STANDARD; PRT; 297 AA.
ID VGLG_HRSV6
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb6256).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; JQ1208; JQ1208.
DR DR InterPro; IPR000925; Glycoprot G.
DR DR Pfam; PF00802; Glycoprotein G; 1.
DR DR Transmembrane; Glycoprotein.
KW
```

```
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

Query Match 87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.5%; Pred. No. 1.2e-80;
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSXNKDQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
DB 1 MSXNKDQRTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
QY 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITTSQTTILASTTP 120
DB 61 FIASANKHVTLLTAIIQDATSQIKNTPTVLTQDPQLGIFSLSNLSITTSQTTILASTTP 120
QY 121 GVKSNLQPTTKTKNTTTTQPSKPTTKQKQKPNKPNNDHFEVFNVPSCISNNP 180
DB 121 SAESTPQSTTVKTKNTTTTQIQPSKPTTKQKQKPNKPNNDHFEVFNVPSCISNNP 180
QY 181 TCWAICKRIENKPKGKTTTKPKPTKTKKDLKPQTTKPKKEVPTTKPTBPTINTTK 240
DB 181 TCWAICKRIENKPKGKTTTKPKPTKTKKDLKPQTTKPKKEVPTTKPTBPTINTTK 240
QY 241 TNIITLLTNTTGNPKLTSQMETFSTSGEGLSPSQVSTTSEHPSQSPSPNTT 296
DB 241 TNIITLLTNTTGNPEYTSQKTLSTSPGPNPSPQVTTTSEYSPSPSPNTT 296

RESULT 5
VGLG_HRSV7 STANDARD; PRT; 297 AA.
ID VGLG_HRSV7
AC P27036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb6614).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11257;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; JQ1209; JQ1209.
DR DR InterPro; IPR000925; Glycoprot G.
DR DR Pfam; PF00802; Glycoprotein G; 1.
DR DR Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
```



```
Db 61 FIASANKVLTATTAIIQDATSOIKNTPTTYLTQNPOLGISFFNLGTTSTQTALITTP 120
QY 121 GVKSNLQPTTKVTKNTTTTQTOPSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
Db 121 SVESIIQSTTVKNTNTTIIQISKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
QY 181 TCWAIKCRIPKPKGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
Db 181 TCWAIKCRIPKPKGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
QY 241 TNITTTLLNTNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPNTTTR 298
Db 241 PNIRTTLLNTSTGNLEHTSQEBTLHSTSEGNLSPSOVSTTSEHPSQSSPNTTTR 298

RESULT 8
VGLG_HRSV2
ID VGLG_HRSV2 STANDARD; PRT; 297 AA.
AC P27021;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb642).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11252;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR; JQ1204; JQ1204.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA; 3745 MW; FC72A7F3A8BFF67C CRC64;

Query Match 84.7%; Score 1307; DB 1; Length 297;
Best Local Similarity 86.5%; Pred. No. 2.6e-78;
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMIISLITAI 60
Db 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMIISLITAI 60
QY 61 FIASANKVLTATTAIIQDATSOIKNTPTTYLTQNPOLGISFFNLGTTSTQTALITTP 120
Db 61 FIASANKVLTATTAIIQDATSOIKNTPTTYLTQNPOLGISFFNLGTTSTQTALITTP 120
QY 121 GVKSNLQPTTKVTKNTTTTQTOPSKPTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
Db 121 GVKLTQSTTVRKNTTTTQQAQPKSTTKORQKNPKPNNDHFEVFNVPVCSICSNPP 180
QY 181 TCWAIKCRIPKPKGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
Db 181 TCWAIKCRIPKPKGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
```

```
QY 241 TNITTTLLNTNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPNTTTR 297
Db 241 TNITTTLSNTARNPELTISQMETFHSTSEGNLSPSOVSTTSEHPSQSSPNTTTR 297

RESULT 9
VGLG_HRSV8
ID VGLG_HRSV8 STANDARD; PRT; 292 AA.
AC P23041;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup B / strain 8/60).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90357765; PubMed=1697126;
RA Sullender W.M., Anderson K., Wertz G.W.;
RT "The respiratory syncytial virus subgroup B attachment glycoprotein:
RT analysis of sequence, expression from a recombinant vector, and
RT evaluation as an immunogen against homologous and heterologous
RT subgroup virus challenge.";
RL Virology 178:195-203 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374595; PubMed=1895391;
RA Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;
RT "Genetic diversity of the attachment protein of subgroup B
RT respiratory syncytial viruses.";
RL J. Virol. 65:5425-5434 (1991).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M55633; AAA47413.1; -.
DR EMBL; M73545; AAA47408.1; -.
DR PIR; A37077; MGNZ60.
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 292 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 292 AA; 32143 MW; 8EC60C85EF0578B5 CRC64;

Query Match 49.7%; Score 767.5; DB 1; Length 292;
Best Local Similarity 55.3%; Pred. No. 2.7e-43;
Matches 162; Conservative 34; Mismatches 94; Indels 3; Gaps 3;

QY 1 MSKNKQORTAKTLEKTDWTLNHLFFISSGLYKLNLSVAQITLSILAMIISLITAI 60
Db 1 MSKNKQORTASTLEKTDWTLNHLFIVISSCLYRLNLSVAQIALSVLAMIISLITAI 60
```



```

Query Match      20.0%; Score 308.5; DB 1; Length 263;
Best Local Similarity 32.5%; Pred. No. 1.5e-13;
Matches 90; Conservative 32; Mismatches 116; Indels 39; Gaps 7;

QY 1 MSKQKQORTAKTLEKWTDTLHLLFISGLYKMLKSAQITLSILAMIISTSLIITAI 60
Db 1 MSNTHHPKFKTLKRAWKASKYFIVGLSCLYKFNKLSLVOTALTTLAMITLSLITAI 60
QY 61 FIASANHKVTLTALIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTOTILASTTP 120
Db 61 YISVGNKAKPTSKPTTQOQLOQNHPTPLTEHNY-----KSTHTSIQSTTL 108
QY 121 GVKNSLQPTTKVTKNTTQTQPSK---PTTKQKQKPPNKP-----NNDHF--FEVF 168
Db 109 SQPNIDTTSSTGYGHTNQTNRKIKSQSTPLATRAPPINPLGSPNPNHQNNSQTL 168
QY 169 NFVPCSCISNPTCWAIC-----KRIPN-----KPKGKTKTKTKTKTKTKKDLK 216
Db 169 PHVPCSTCEGNPACSPLCQIELERAPSSAPTITLKKAPKTKTKTKTKTKTIYHRTSPEAK 228
QY 217 POTTKEVPTTKPTTEPTINTKNTITNTLLTNNTT 253
Db 229 LQTKKIMVTPQOQILSSP---EHQTNQSTTQISQHTS 262

RESULT 12
VGLG BRSV4
ID VGLG BRSV4 STANDARD; PRT; 263 AA.
AC Q10684;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain 4642) (BRS).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=82820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97288324; PubMed=9143302;
RA Furze J., Roberts S., Wertz G., Taylor G.;
RT "Antigenically distinct G glycoproteins of BRSV strains share a high
RL degree of genetic homogeneity.";
RL Virology 231:48-58(1997).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
CC HRS VIRUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y08718; CAAG9968.1; --
CC HSSP; P22261; 1BRV.
CC InterPro; IPR000925; Glycoprot.g.
CC Pfam; PF00802; Glycoprotein G; 1.
CC Transmembrane; Glycoprotein.
CC DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 38 66 POTENTIAL.
CC DOMAIN 67 263 EXTRACELLULAR (POTENTIAL).
CC DISULFID 173 186 BY SIMILARITY.
CC DISULFID 176 182 BY SIMILARITY.
CC CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 263 AA; 28976 MW; D592D79EF8EE9535 CRC64;

QY 1 MSKQKQORTAKTLEKWTDTLHLLFISGLYKMLKSAQITLSILAMIISTSLIITAI 60
Db 1 MSNTHHPKFKTLKRAWKASKYFIVGLSCLYKFNKLSLVOTALTTLAMITLSLITAI 60
QY 61 FIASANHKVTLTALIQDATSQIKNTTPTVLTQDPQLGISFNSLSEITSTOTILASTTP 120
Db 61 YISVGNKAKPTSKPTTQOQLOQNHPTPLTEHNY-----KSTHTSIQSTTL 108
QY 121 GVKNSLQPTTKVTKNTTQTQPSK---PTTKQKQKPPNKP-----NNDHF--FEVF 168
Db 109 SQPNIDTTSSTGYGHTNQTNRKIKSQSTPLATRAPPINPLGSPNPNHQNNSQTL 168
QY 169 NFVPCSCISNPTCWAIC-----KRIPN-----KPKGKTKTKTKTKTKTKKDLK 216
Db 169 PHVPCSTCEGNPACSPLCQIELERAPSSAPTITLKKAPKTKTKTKTKTKTIYHRTSPEAK 228
QY 217 POTTKEVPTTKPTTEPTINTKNTITNTLLTNNTT 253
Db 229 LQTKKIMVTPQOQILSSP---EHQTNQSTTQISQHTS 262

RESULT 13
VGLG BRSVC
ID VGLG BRSVC STANDARD; PRT; 257 AA.
AC P22261;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain Copenhagen) (BRS).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91012801; PubMed=2214024;
RA Lerch R.A., Anderson K., Wertz G.W.;
RT "Nucleotide sequence analysis and expression from recombinant vectors
RT demonstrate that the attachment protein G of bovine respiratory
RT syncytial virus is distinct from that of human respiratory syncytial
RT virus.";
RL J. Virol. 64:5559-5569(1990).
RN [2]
RP STRUCTURE BY NMR OF 158-189.
RC STRAIN=391-2;
RX MEDLINE=97098087; PubMed=8942628;
RA Doreleijers J.F., Langedijk J.P.M., Haard K., Boelens R.,
RA Rullmann J.A., Schaap W.M., van Oirschot J.T., Kaptein R.;
RT "Solution structure of the immunodominant region of protein G of
RT bovine respiratory syncytial virus.";
RL Biochemistry 35:14684-14688(1996).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
CC HRS VIRUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58307; AAA42810.1; --

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:33:35 ; Search time 28.6755 Seconds  
(without alignments)  
999.400 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKQRTAKTLEKWTDL.....VSTTSHPSQSPSPNTTQ 298

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	99.1	298	1 MGNZRL	major surface glyc
2	1418	91.8	298	1 MGNZ	major surface glyc
3	1343	87.0	297	2 JQ1205	attachment protein
4	1343	87.0	297	2 JQ1208	attachment protein
5	1333	86.3	298	2 JQ5680	G protein - Human
6	1328	86.0	297	2 JQ1209	attachment protein
7	1324	85.8	297	2 JQ1206	attachment protein
8	1313	85.0	298	2 JQ1207	attachment protein
9	1307	84.7	297	2 JQ1204	attachment protein
10	767.5	49.7	292	1 MGNZ60	major surface glyc
11	761	49.3	292	1 MGNZ18	major surface glyc
12	425	27.5	467	1 VHNZ	nucleocapsid prote
13	299	19.4	250	2 P00768	glycoprotein G - b
14	299	19.4	257	1 MGNZR	major surface glyc
15	295.5	19.1	263	2 JQ2284	glycoprotein G - b
16	284	18.4	263	2 JQ2288	glycoprotein G - o
17	280	18.1	248	2 P00769	glycoprotein G - b
18	272	17.6	263	2 A48732	attachment glycopr
19	208	13.5	307	1 GSPF3	salivary glue prot
20	194	12.6	662	2 A45155	mucin FIM-C.1 - Af
21	193	12.5	3020	2 A43932	mucin 2 precursor,
22	191	12.4	851	2 T22696	hypothetical prote
23	191	12.4	1161	2 S57180	probable membrane
24	190	12.3	322	2 A53715	apomucin precursor
25	188	12.2	1832	2 T31113	mucin-like glycopr
26	185.5	12.0	379	2 S50125	larval glue protei
27	183	11.9	217	2 S01358	salivary glue prot
28	183	11.9	354	2 T46740	microfilarial shea
29	177	11.5	770	2 T22808	hypothetical prote

30 175 11.3 825 2 T29634  
31 172.5 11.2 327 2 S20074  
32 172.5 11.2 1777 2 T34369  
33 171.5 11.1 216 2 T51920  
34 170.5 11.0 279 2 S53363  
35 170 11.0 3570 2 T45025  
36 169.5 11.0 1367 1 S48478  
37 168 10.9 402 2 E86185  
38 167.5 10.8 294 2 A37232  
39 167.5 10.8 1630 2 A53577  
40 164.5 10.7 2476 2 T34022  
41 163 10.6 371 2 S20075  
42 163 10.6 477 2 S53362  
43 160.5 10.4 393 2 S62335  
44 160.5 10.4 866 2 T45462  
45 160.5 10.4 867 2 T45463

hypothetical prote  
promastigote surfa  
hypothetical prote  
mucin - rhesus mac  
mucin SAC (clone J  
mucin MUC5B, trach  
glucan 1,4-alpha-g  
hypothetical prote  
mucin, tracheal (A  
ascites sialoglyco  
zonadhesin - pig  
promastigote surfa  
mucin SAC (clone J  
171-7 protein - fr  
membrane glycoprot  
membrane glycoprot

## ALIGNMENTS

### RESULT 1

MGNZRL

major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C:Accession: A32703, S12279

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: A32703; UID:87289657; PMID:2441388

A:Accession: A32703

A:Molecule type: mRNA

A:Residues: 1-298 <OH>

A:Cross-references: GB:M17212; NID:G333940; PIDN:AAA47411.1; PID:G333941

R:Garcia-Barreno, B.; Portela, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.

EMBO J. 9, 4181-4187, 1990

A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizati

A:Reference number: S12279; UID:91065351; PMID:2249671

A:Accession: S12279

A:Molecule type: mRNA

A:Residues: 1-298 <GAR>

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:85,103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 99.1%; Score 1530; DB 1; Length 298;

Best Local Similarity 99.0%; Pred. No. 2.3e-92;

Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKWTDLNHLFPISGLYKLNKSVQAQITLSILAMITSLITAIL 60

DB 1 MSKNKQRTAKTLEKWTDLNHLFPISGLYKLNKSVQAQITLSILAMITSLITAIL 60

QY 61 FIASANKVLTALIQDTSQIKNTTPTLTQDPLGIFSNLSITSTQTTILASTTP 120

DB 61 FIASANKVLTALIQDTSQIKNTTPTLTQDPLGIFSNLSITSTQTTILASTTP 120

QY 121 GYKSNLOPTVTKNTTQTOTQSPKPTTKQKPNKPNNDHFVEVNFVPCISNNP 180

DB 121 GYKSNLOPTVTKNTTQTOTQSPKPTTKQKPNKPNNDHFVEVNFVPCISNNP 180

QY 181 TCWAICKRIPNKKPKKTKTTTKPTTKPTTKTKDLPQTTPKPKVPTTKPTEETINTTK 240

DB 181 TCWAICKRIPNKKPKKTKTTTKPTTKPTTKTKDLPQTTPKPKVPTTKPTEETINTTK 240

QY 241 TMTTTLTNNTGPNKLTQSMETPHSTSSSEGNLSPSQVSTTSHPSQSPSPNTTQ 298

DB 241 TMTTTLTNNTGPNKLTQSMETPHSTSSSEGNLSPSQVSTTSHPSQSPSPNTTQ 298

### RESULT 2

## MGZ

major surface glycoprotein G - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C>Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 24-Sep-1999  
C:Accession: A94048; A93599; A04039  
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
A:Reference number: A94048; MUID:85216636; PMID:3858865  
A:Accession: A94048  
A:Molecule type: mRNA  
A:Residues: 1-298 <WER>  
A:Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:M11217; GB:7.1; PID:G333932  
A>Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
R:Satake, M.; Colligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S.  
Nucleic Acids Res. 13, 7795-7812, 1985  
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
A:Reference number: A93599; MUID:86067198; PMID:4069997  
A:Accession: A93599  
A:Molecule type: mRNA  
A:Residues: 1-298 <SAT>  
A:Cross-references: GB:X03149; NID:G60997; PIDN:CAA26928.1; PID:G60998  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:38-66/Domain: transmembrane #status predicted <TNM>  
F:85,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 1418; DB 1; Length 298;  
Best Local Similarity 93.3%; Pred. No. 4.2e-85;  
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
Qy 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Db 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Qy 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Db 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Db 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Qy 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Db 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 298  
Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 298

## RESULT 3

JQ1205  
attachment protein - human respiratory syncytial virus (strain RSB1734)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1205  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1205  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;  
Best Local Similarity 88.6%; Pred. No. 3.1e-80;  
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;  
Qy 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Db 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Qy 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Db 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Db 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Qy 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Db 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 297  
Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 297

## RESULT 4

JQ1208  
attachment protein - human respiratory syncytial virus (strain RSB6256)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1208  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1208  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;  
Best Local Similarity 88.5%; Pred. No. 3.1e-80;  
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;  
Qy 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Db 1 MSKNKDQRTAKTLKKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
Qy 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Db 61 FIASANKHVLTITAIQDATSQIKNTTPTVLTQDPOLGISFNSLSEITSTQTTILASTTP 120  
Qy 121 GVKSNILOPTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Db 121 SAESTQSTTVTKNTTQTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
Qy 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Db 181 TCWAICKRIIPNKKPKGKTTTKPTTKPTTKKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
Qy 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 296  
Db 241 TNITTTLLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTTR 296

## RESULT 5

JC5680

G protein - Human respiratory syncytial virus  
 C:Species: human respiratory syncytial virus  
 C>Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
 C:Accession: J05680  
 R:Geng, X.; Wang, Z.; Qian, Y.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
 Chinese J. Virol. 12, 317-322, 1996  
 A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain is  
 A:Reference number: J05680  
 A:Accession: J05680  
 A:Molecule type: mRNA  
 A:Residues: 1-298 <GEN>  
 A:Experimental source: strain B79  
 A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 F:1-38/Domain: intracellular #status predicted <INT>  
 F:39-66/Domain: transmembrane #status predicted <TM>  
 F:67-298/Domain: extracellular #status predicted <EXC>

Query Match 86.3%; Score 1333; DB 2; Length 298;  
 Best Local Similarity 87.9%; Pred. No. 1.4e-79;  
 Matches 262; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPOSTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240

QY 241 TNITTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296

RESULT 6  
 J01209  
 attachment protein - human respiratory syncytial virus (strain RSB6614)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01209  
 R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A:Reference number: J01209; MUID:91374005; PMID:1895054  
 A:Accession: J01209  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <CAN>  
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract in  
 children and adults.  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 86.0%; Score 1328; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 2.9e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPOSTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240

QY 241 TNITTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296

RESULT 7  
 J01206  
 attachment protein - human respiratory syncytial virus (strain RSB5857)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01206  
 R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
 A:Reference number: J01204; MUID:91374005; PMID:1895054  
 A:Accession: J01206  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <CAN>  
 C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract  
 children and adults.  
 C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C:Keywords: glycoprotein; transmembrane protein  
 F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 85.8%; Score 1324; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 5.3e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180  
 DB 121 SAESTPOSTTVTKNTTNTTQPSKPTTKORQKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKKTITKPTKPTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTK 240

QY 241 TNITTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296  
 DB 241 TNIRTTLLTNTTGNPKLTSQMETPHSTSSBGNLSPSOVSTTSBHPSPSPSPNTT 296

RESULT 8  
 J01207  
 attachment protein - human respiratory syncytial virus (strain RSB6190)  
 N:Alternate names: G protein  
 C:Species: human respiratory syncytial virus  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C:Accession: J01207  
 R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A:Title: Identification of variable domains of the attachment (G) protein of subgroup  
 A:Reference number: J01204; MUID:91374005; PMID:1895054  
 A:Accession: J01207  
 A:Molecule type: mRNA

Query Match 86.0%; Score 1328; DB 2; Length 297;  
 Best Local Similarity 87.8%; Pred. No. 2.9e-79;  
 Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60  
 DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKMLKSAQITLSILAMIISTSLIIITAI 60

QY 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120  
 DB 61 FIASANKHVLTITAIQDATSQIKNTTPTTYLTQDQPGISFNSLSEITSTQTTILASTTP 120

A;Residues: 1-298 <CAN>  
A;Note: the authors translated the codon ACC for residue 4 as Asn and AGC for residue 22 as Arg. The authors did not indicate the position of the stop codon in the sequence.  
C;Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract illness in children and adults.  
C;Superfamily: respiratory syncytial virus major surface glycoprotein G  
C;Keywords: glycoprotein; transmembrane protein  
F;85,103,135,237,250,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.0%; Score 1313; DB 2; Length 298;  
Best Local Similarity 86.9%; Pred. No. 2.7e-78;  
Matches 259; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDLNHLPLFTSSGLYKLNLSVAQITLSILAMIISTSLIITAI 60  
Db 1 MSKTKQORTAKTLEKWTDLNHLPLFTSSGLYKLNLSVAQITLSILAMIISTSLIAAII 60  
QY 61 FIASANHKVLTLLAIQDATSQIKNTTPTVLTODPOLGISFNSLSEITSTOTTILASTTP 120  
Db 61 FIASANNKVLTLLAIQDATSQIKNTTPTVLTQNPQLGISFNLGTTSTOTTAILALTTP 120  
QY 121 GVKSNLOPTTVTKNTTQTTPQSPKPTTKQRQKNPKPNNDHFEVFNVPVPCISCSNPP 180  
Db 121 SVESILQSTTVTKNTTQTTPQSPKPTTKQRQKNPKPNNDHFEVFNVPVPCISCSNPP 180  
QY 181 TCWAIKCRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKPEVTTKTEPTTINTYK 240  
Db 181 TCWAIKCRIPSKPGKTTTKPKTKPTTKKDLKPQTKPKPEAPSTKPKTEPTTINIK 240  
QY 241 TNTITTTLLTNNTGNPKLTSQMETFTSTSEGNLSPSOVSTTSEHPSPQSPSPNTTQ 298  
Db 241 PNIRTTLLTNSTTGNLEHTSQEFTLSTSEGNLSPSOVTTSEYLSQTPSPSNIQT 298

RESULT 9  
JQ1204  
Attachment protein - human respiratory syncytial virus (strain RSB642)  
N;Alternate names: G protein  
C;Species: human respiratory syncytial virus  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
R;Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2091-2096, 1991  
A;Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A;Reference number: JQ1204; MUID:91374005; PMID:1895054  
A;Accession: JQ1204  
A;Molecule type: mRNA  
A;Residues: 1-297 <CAN>  
C;Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract illness in children and adults.  
C;Superfamily: respiratory syncytial virus major surface glycoprotein G  
C;Keywords: glycoprotein; transmembrane protein  
F;85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.7%; Score 1307; DB 2; Length 297;  
Best Local Similarity 86.5%; Pred. No. 6.7e-78;  
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDLNHLPLFTSSGLYKLNLSVAQITLSILAMIISTSLIITAI 60  
Db 1 MSKNKQORTAKTLERTWDTLNHLPLFTSSGLYKLNLSVAQITLSILAMIISTSLIAAII 60  
QY 61 FIASANHKVLTLLAIQDATSQIKNTTPTVLTODPOLGISFNSLSEITSTOTTILASTTP 120  
Db 61 FIASANHRVTSITLLAIQDATNQIKNTTPTVLTQNPQLGISPNPSEITSLITILDPTTP 120  
QY 121 GVKSNLOPTTVTKNTTQTTPQSPKPTTKQRQKNPKPNNDHFEVFNVPVPCISCSNPP 180  
Db 121 GVXLTLQSTTVTKNTTQTTPQSPKPTTKQRQKNPKPNNDHFEVFNVPVPCISCSNPP 180  
QY 181 TCWAIKCRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKPEVTTKTEPTTINTYK 240  
Db 181 TCWAIKCRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKPEVTTKTEPTTINTYK 240

Best Local Similarity 54.9%; Pred. No. 1.7e-42;  
Matches 158; Conservative 35; Mismatches 93; Indels 2; Gaps 2;

QY 1 MSKNKQRTAKTEKTDWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
DB 1 MSKHQRTAKTEKTDWTLNHLIVISSCLYRLNLSIAQIALSVLAMIISTLSIIITAI 60

QY 61 FIASANKHVLTITTAIQDQTSQIKNTTPTVLTQDQGLGIFSNLSIETSTTTILASTTP 120  
DB 61 FIISANKHVLTITTVQTIKNHTEKNISTVLTQVPPERVNSKQPTTTSPIHTNSATISP 120

QY 121 GVKSNLOPTVTKNTTITOTOPSKPTTKORQKNPKNNDFHFEVFNVPVCSICSNPP 180  
DB 121 NTKSETHHTAQTKRITTSQTNKPSKSRNPPKPKDDHFEVFNVPVCSICGNQ 180

QY 181 TCMAICKRIIPNKKPKKTKPTTKPTKTT-KDGLKQPTTKPKVEPTTKPTKEPTINTT 239  
DB 181 LKSIKCTIPSNKPKKPTKPTKPTKTTKNTKREDKTPAKMPKKEIITNPAPKPKTKTT 240

QY 240 KTNITITLLTNNTGNKLSQMETFSTSEGNLSPSVSTTSEHPS 287  
DB 241 ERDTSISQSTVLDITTPKTYIQOQSLHSTTSENTPTSTQIPTASE-PS 287

RESULT 12  
VNZ  
nucleocapsid protein (version 2) - human respiratory syncytial virus  
C:Species: human respiratory syncytial virus  
C>Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999  
C:Accession: A04026  
R:Elango, N.; Venkatesan, S.  
Nucleic Acids Res 11, 5941-5951, 1983  
A>Title: Amino acid sequence of respiratory syncytial virus capsid protein.  
A:Reference number: A04026; MUID:83299261; PMID:6310521  
A:Accession: A04026  
A:Molecule type: mRNA  
A:Residues: 1-467 <ELA>  
A:Cross-references: GB:X00001; NID:G61215; PIDN:CRA24906.1; PID:G61216  
C:Genetics:  
A:Gene: N  
C:Superfamily: respiratory syncytial virus nucleocapsid protein  
C:Keywords: nucleocapsid

Query Match 27.5%; Score 425; DB 1; Length 467;  
Best Local Similarity 90.2%; Pred. No. 1.7e-20;  
Matches 83; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 207 TFKTKKDLKQPTTKPKVEPTTKPTKEPTINTTNTTLLTNNTGNKLSQMETFTH 266  
DB 376 TLKTKKDLKQPTTKPKVEPTTKPTKEPTINTTNTTLLTNNTGNPELTSQMETFTH 435

QY 267 STSSEGNLSPSVSTTSEHPSQSPSPNTTQ 298  
DB 436 STSSEGNLSPSVSTTSEHPSQSPSPNTTQ 467

RESULT 13  
PQ0768  
glycoprotein G - bovine respiratory syncytial virus (isolate FS-1) (fragment)  
C:Species: bovine respiratory syncytial virus  
C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999  
C:Accession: PQ0768  
R:Malipeddi, S.K.; Samal, S.K.  
J. Gen. Virol. 74, 2001-2004, 1993  
A>Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial virus.  
A:Reference number: J02284; MUID:93389461; PMID:8376974  
A:Accession: PQ0768  
A:Molecule type: mRNA  
A:Residues: 1-250 <MAL>  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:1-31/Domain: intracellular #status predicted <INT>  
F:32-59/Domain: transmembrane #status predicted <TM>

F:60-250/Domain: extracellular #status predicted <EXT>  
F:78,156,226,241,244/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 2; Length 250;  
Best Local Similarity 30.3%; Pred. No. 1.3e-12;  
Matches 91; Conservative 35; Mismatches 104; Indels 70; Gaps 9;

QY 11 KTLKNTWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAIPIASANKHV 70  
DB 4 KTLKNAKASYFIVGLSCLYKFNLSLVQVLTSLAMITLTSIVITAIIVISGNAKAK 63

QY 71 LTTALIQDQTSQIKNTTPTVLTQDQGLGIFSNLSIETSTTTILASTTPGVKSNLOPTT 130  
DB 64 PTKPTTIQOTQOPQNHSTSPFFTENNY-----KSTHTSIQSTTLLSOLINIDTTR 111

QY 131 VKTKNTTITOTQPSK-----PTTKQ-----RONKPPKPNNDHFEVFNVPVCSICSN 178  
DB 112 GTTYGHSTDETSQSKIKSQSTLPTTRKPPINPSSNPENHODHNSQTLPIVPCSTCEG 171

QY 179 NPTCWAICKRIIPNKKPKKTKPTTKPTTKTKDILKQPTTKPKVEPTTKPTKEPTINTT 238  
DB 172 NLACLSLCQVGR-----APSRAPTI-TLAKTPKPTTK-----KP----- 207

QY 239 TKNTITITLLTNNTGNKLSQMETFSTSEGNLSPSVSTTSEHPSQSPSPNTTQ 298  
DB 208 ----IKTII-----HRTSPEAKLPKN-NTAAPQOGLSSPENTNQ 245

RESULT 14  
MGNZBR  
major surface glycoprotein G - bovine respiratory syncytial virus (strain 391-2)  
N:Alternate names: attachment glycoprotein G  
C:Species: bovine respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: A36408  
R:Lerch, R.A.; Anderson, K.; Wertz, G.W.  
J. Virol. 64, 5559-5569, 1990  
A>Title: Nucleotide sequence analysis and expression from recombinant vectors demonstr  
Y syncytial virus.  
A:Reference number: A36408; MUID:91012801; PMID:2214024  
A:Accession: A36408  
A:Molecule type: mRNA  
A:Residues: 1-257 <LER>  
A:Cross-references: GB:M58307; NID:G210830; PIDN:AAA42810.1; PID:G210831  
C:Genetics:  
A:Gene: G  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:45-62/Domain: transmembrane #status predicted <TM>  
F:73,85,127,149,233,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 1; Length 257;  
Best Local Similarity 31.6%; Pred. No. 1.3e-12;  
Matches 92; Conservative 33; Mismatches 120; Indels 46; Gaps 8;

QY 1 MSKNKQRTAKTEKTDWTLNHLIFISSGLYKLNKLSVAQITLSILAMIISTLSIIITAI 60  
DB 1 MSNHTHLKFKTLKRAWKASKYFIVGLSCLYKFNLSLVQVLTSLAMITLTSIVITAI 60

QY 61 FIASANKHVLTITTAIQDQTSQIKNTTPTVLTQDQGLGIFSNLSIETSTTTILASTTP 120  
DB 61 YISVGNARAKPTKPTTIQOTQOPQNHSTSPFFTENNY-----KSTHTSIQSTTLL 108

QY 121 GVKSNLOPTVTKNTTITOTQPSK-----PTTKQKNPKPNNDHFEVFNVPVCSICSN 168  
DB 109 SOLINIDTTRGITTYGHSTNEQNKRKIQOSTLPATRPPIINPSSIPENHODHNNFTL 168

QY 169 NFVPCSICSNPTCWAICKRIIPNKKPKKTKPTTKPTTKTKDILKQPTTKPKVEPTT 228  
DB 169 PVVPCSTCEGNLACLSCH-----IETRAPSRAPTI-TLAKTPKPTTK----- 212

QY 229 KPTKEPTINTTNTTLLTNNTGNKLSQMETFSTSEGNLSPSVSTTSEHPSQSP 279

Db 213 KPTK--TTIHRSTPETHKLPKNNATP-----QQGILSSTEHHNQSTTQI 257

## RESULT 15

JQ2284

glycoprotein G - bovine respiratory syncytial virus (isolate A51908)

C;Species: bovine respiratory syncytial virus

C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-May-1999

C;Accession: JQ2284

R;Mallipeddi, S.K.; Samal, S.K.

J. Gen. Virol. 74, 2001-2004, 1993

A;Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial v

A;Reference number: JQ2284; MUID:93389461; PMID:8376974

A;Accession: JQ2284

A;Molecule type: mRNA

A;Residues: 1-263 &lt;MAL&gt;

A;Experimental source: isolate A51908

C;Superfamily: respiratory syncytial virus major surface glycoprotein G

C;Keywords: glycoprotein; transmembrane protein

F;1-38/Domain: intracellular #status predicted &lt;CYT&gt;

F;39-66/Domain: transmembrane #status predicted &lt;TM&gt;

F;67-263/Domain: extracellular #status predicted &lt;EXT&gt;

F;127,163,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

	Query Match	19.1%	Score 295.5;	DB 2;	Length 263;
	Best Local Similarity	31.7%	Pred. No. 2.3e-12;		
	Matches	86;	Conservative	35;	Mismatches 123; Indels 27; Gaps 7;
Qy	1	MSKNKQRTATLSEKTDWTLNHLFISSGLYKLNKSKVAQITLSILAMIISTSLITAIL	60		
Db	1	MSNHTHPKFKTLKRAMKASKYFIVGLSCLYFNLSIVQTALTSLAMITLSLVITAIL	60		
Qy	61	FIASANKHVTLLTALIQDATSQIKNTTPTYLTDPPQLGISFNSLSEITSTQTTILASTTP	120		
Db	61	YISVGNKAKETSPTTQQTQOPQNHPTLLPTEHNHKSHTSTSTQTTLSQPPNI--DTTS	118		
Qy	121	GVK-----SNLQPTTVTKTNT--TTTQPSKPTTKQRONKPPKPNNDHFPEVFNVP	174		
Db	119	GTTYGHPINRTQNRKIKSQSTPLATKLPINPL-----ESNPPENHQDHNSQTLPHVPCS	174		
Qy	175	ICSNNPICWAIC----KRIPN-----KPGKKTTTKPKTKTKKOLKPKQTKP	222		
Db	175	TCEGNPACSPICQIGLERAPSRAPTITLKAKPKTKTKTKTKTIYHRTSPEAKLQTKN	234		
Qy	223	KEVPTTKPTERTPTINTTKNTITTTLLTNNTT	253		
Db	235	TATPQQGILSP---EHGTNQSTTQISQHTS	262		

Search completed: October 29, 2003, 17:41:49

Job time : 31.6755 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:41:00 ; Search time 52.2906 Seconds  
(without alignments)  
975.349 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNQDQRTAKLEKTDWL.....VSTTSEHPSPQSPPPNTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 642050 seqs, 171146064 residues

Total number of hits satisfying chosen parameters: 642050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	193	12.5	5179	US-09-922-217-1068
2	193	12.5	5179	Sequence 1068, Ap
3	193	12.5	5179	Sequence 1068, Ap
4	177.5	11.5	800	US-10-025-380-1068
5	169.5	11.0	1367	US-10-029-386-32198
6	169	10.9	36	US-09-801-368-108
7	162.5	10.5	288	US-10-351-641-871
8	162.5	10.5	288	Sequence 341, App
9	162.5	10.5	288	US-09-216-393-341
10	162.5	10.5	288	Sequence 344, App
11	162	10.5	288	US-10-321-856-341
12	162	10.5	941	US-10-124-557-14
13	162	10.5	1022	US-10-124-557-84
14	162	10.5	1038	US-10-124-557-74
15	162	10.5	1049	US-10-124-557-58
16	162	10.5	1140	US-10-124-557-104

16	162	10.5	1270	14	US-10-124-557-44	Sequence 44, Appl
17	162	10.5	1311	14	US-10-124-557-42	Sequence 42, Appl
18	162	10.5	1313	14	US-10-124-557-142	Sequence 142, App
19	162	10.5	1314	14	US-10-124-557-50	Sequence 50, Appl
20	162	10.5	1320	14	US-10-124-557-46	Sequence 46, Appl
21	162	10.5	1320	14	US-10-124-557-60	Sequence 60, Appl
22	162	10.5	1354	14	US-10-124-557-48	Sequence 48, Appl
23	162	10.5	1361	14	US-10-124-557-40	Sequence 40, Appl
24	162	10.5	1363	14	US-10-124-557-52	Sequence 52, Appl
25	162	10.5	1404	14	US-10-124-557-2	Sequence 2, Appli
26	162	10.5	1404	14	US-10-124-557-62	Sequence 62, Appl
27	155	10.0	1260	15	US-10-245-802-8	Sequence 8, Appli
28	154	10.0	528	12	US-09-840-746-20	Sequence 20, Appl
29	152.5	9.9	538	12	US-10-038-694-3	Sequence 3, Appli
30	150.5	9.7	995	11	US-09-984-130-48	Sequence 48, Appl
31	150.5	9.7	995	12	US-09-836-353A-48	Sequence 48, Appl
32	145	9.4	1075	10	US-09-801-368-110	Sequence 110, App
33	144.5	9.4	1322	10	US-09-801-368-114	Sequence 114, App
34	144.5	9.4	2828	10	US-09-905-129-21	Sequence 21, Appl
35	144.5	9.4	2828	10	US-09-905-129-21	Sequence 21, Appl
36	144.5	9.4	2828	12	US-10-301-822-49	Sequence 49, Appl
37	144.5	9.4	2828	12	US-10-032-189-126	Sequence 126, App
38	144.5	9.4	2828	15	US-10-176-847-54	Sequence 54, Appl
39	144.5	9.4	2828	15	US-10-177-293-110	Sequence 110, App
40	144	9.3	449	12	US-10-137-870-224	Sequence 224, App
41	144	9.3	449	12	US-10-140-018-224	Sequence 224, App
42	144	9.3	449	12	US-10-140-021-224	Sequence 224, App
43	144	9.3	449	12	US-10-140-274-224	Sequence 224, App
44	144	9.3	449	12	US-10-140-471-224	Sequence 224, App
45	144	9.3	449	12	US-10-140-807-224	Sequence 224, App

## ALIGNMENTS

### RESULT 1

US-09-922-217-1068  
; Sequence 1068, Application US/09922217  
; Patent No. US20020076414A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1068  
; LENGTH: 5179  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-217-1068

Query Match 12.5%; Score 193; DB 9; Length 5179;  
Best local Similarity 30.6%; Pred. No. 2.3e-06;  
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;  
71 LTTAIIQDA--TSQIKNTTP----TYLTQDQLGTSFNSLSEIT--SQTTLASTTPG 121  
1504 MTTPIPPASTTTLPTTTPSPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPS 1563

RESULT 3  
US-10-025-380-1068  
; Sequence 1068, Application US/10025380  
; Publication No. US20020182191A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.

RESULT 4  
US-10-029-386-32198  
; Sequence 32198, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Randk, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32198  
; LENGTH: 800  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC007663.28  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3  
US-10-029-386-32198

Query Match 11.5%; Score 177.5; DB 12; Length 800;  
Best Local Similarity 23.7%; Pred. No. 3.9e-06;  
Matches 69; Conservative 39; Mismatches 148; Indels 35; Gaps 4;

QY 9 TATLEKTDWTLNHLFISSGLYKLNKVAQITLSILAMISTSLIITAIIFASANK 68  
DB 325 TTTTITT 380

QY 69 VTLLTAIIOATSOIKNTTPTYLTOPQLGIFSNSLSEITSOITTLASTTPGVKSNLQP 128  
DB 381 ITITT 440

QY 129 TTVKTKNTTITQOPSKPTTKQKQKPPKPNNDHFVEFVPCSCSNPTCWAICKR 188  
DB 441 TTTITSTT 475

QY 189 IPNKKPGKTKTKTKPTTKTKKOLKPOTKPKVEPTT---KPTSEPTINTKINI-- 243  
DB 476 TTTTITTI 535

QY 244 -TTLLNNTGNPKLTSOMETHSTSEGNLSPSQVSTSEHPSPQSPSP 293  
DB 536 TTTTITT 586

## RESULT 5

US-09-801-368-108  
; Sequence 108, Application US/09801368  
; Patent No. US20020128250A1

; GENERAL INFORMATION:  
; APPLICANT: Busby, Robert

; APPLICANT: Cali, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: No. US20020128250A1man, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir

; APPLICANT: Silva, Jeff

; APPLICANT: Summers, Eric

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147

; CURRENT APPLICATION NUMBER: US/09/801.368

; CURRENT FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 09/487,558

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US 60/160,587

; PRIOR FILING DATE: 1999-10-20

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 108

; LENGTH: 1367

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-801-368-108

Query Match 11.0%; Score 169.5; DB 10; Length 1367;  
Best Local Similarity 26.2%; Pred. No. 3.5e-05;  
Matches 62; Conservative 36; Mismatches 94; Indels 45; Gaps 8;

QY 84 KNTTPTLTQPOLGIFSNSLSEITSOITTLASTTPGV--KSNLQPTTKTKNTT--T 139  
DB 209 KSSITTSSTT 268

QY 140 QTQPSKPTTKQKQKPPKPNNDHFVEFVPCSCSNPTCWAICKRIPNKKPGKKT 199  
DB 269 KEKTPPTTTCTKKEKTPPHD-----TTPCT-----KKKTTTSKTCCTKKT 310  
QY 200 TKPTKPTKTKTKKDLKP-----QTTKPKVEPTTKPTEB-----PTINTTKINITTL 247  
DB 311 TTPPTTSSSTTSSSAPVPTPSSSTTSSSAPVTSSTTSSSAPVPTPSSSTTSSSAP 370  
QY 248 LTNNNTGNPKLTSOMETHSTSEGNLSPSQVSTTSEHPSPQ-----SSPNTT 296  
DB 371 VTSSTT---ESSAPVTSSTTSSSAPVPTPSSSTTSSSAPVTSSTTSSSAPVTS 424

## RESULT 6

US-10-351-641-871

; Sequence 871, Application US/10351641

; Publication No. US20030186874A1

; GENERAL INFORMATION:

; APPLICANT: Bardey, S.

; APPLICANT: Guthrie, K.

; APPLICANT: Merutka, G.

; APPLICANT: Anwer, M.

; APPLICANT: Lambert, D.

; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC

; FILE REFERENCE: 7872-100

; CURRENT APPLICATION NUMBER: US/10/351.641

; CURRENT FILING DATE: 2003-01-24

; PRIOR APPLICATION NUMBER: 09/350.641

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: 09/315.304

; PRIOR FILING DATE: 1999-05-20

; PRIOR APPLICATION NUMBER: 09/082.279

; PRIOR FILING DATE: 1998-05-20

; NUMBER OF SEQ ID NOS: 1757

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 871

; LENGTH: 36

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Core polypeptide

US-10-351-641-871

Query Match 10.9%; Score 169; DB 12; Length 36;

Best Local Similarity 94.4%; Pred. No. 3.5e-07;

Matches 34; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 AKTLEKTDWTLNHLFISSGLYKLNKVAQITLSI 45

DB 1 AKTLEKTDWTLNHLFISSGLYKLNKVAQITLSI 36

## RESULT 7

US-09-216-393-341

; Sequence 341, Application US/09216393

; Patent No. US2001001447A1

; GENERAL INFORMATION:

; APPLICANT: Milhausen, Michael James

; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND

; FILE REFERENCE: TX-1-C2

; CURRENT APPLICATION NUMBER: US/09/216.393

; CURRENT FILING DATE: 1998-12-18

; EARLIER APPLICATION NUMBER: 08/994,825

; EARLIER FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 364

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 341

; LENGTH: 288

; TYPE: PRT

; ORGANISM: Toxoplasma gondii

US-09-216-393-341

Query Match 10.5%; Score 162.5; DB 9; Length 288;  
Best Local Similarity 23.1%; Pred. No. 1.8e-05;  
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;  
  
QY 39 AQTLSILAMISTSLIITAIIFIASANHKVLTITAIQDATSOI----- 83  
DB 76 AALSTDAIDRVSQFDVSLDLVIREAAQKFDLLGLRLITDIASGIGEGAMALMGEEAAFI 135  
QY 84 -----KNTPTYLQDPQLGIFSNLSBITSQTTILASTTPGVKSNLQPTTVTKN 135  
DB 136 RPRSRKRGKTTT-----TSSSTSTSTTTTSTTTT----- 181  
QY 136 TTTTQPSKPTTKQKQKPNKPNDFHFEVFNPCISNNPTCWAICKRIPNKKPG 195  
DB 182 TTTTITTTTPTT----- 194  
QY 196 KTTTTPKPTKPTTKKOLKQPTTKPEVPTTKPTBEPTINTKTNTITLLTNNTGN 255  
DB 195 -----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTSTTTTSTTTT 249  
QY 256 PKLTSQMETFHSSEGNLSPQSVSTTSEHPSQSPSPNTT 296  
DB 250 PTTT-----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTTTTSTTTT 284  
  
RESULT 8  
US-09-216-393-344  
; Sequence 344, Application US/09216393  
; Patent No. US20010014447A1  
; GENERAL INFORMATION:  
; APPLICANT: Milhausen, Michael James  
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND  
; FILE REFERENCE: TX-1-C2  
; CURRENT APPLICATION NUMBER: US/09/216.393  
; EARLIER FILING DATE: 1998-12-18  
; EARLIER FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 344  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Toxoplasma gondii  
US-09-216-393-344

Query Match 10.5%; Score 162.5; DB 9; Length 288;  
Best Local Similarity 23.1%; Pred. No. 1.8e-05;  
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;  
  
QY 39 AQTLSILAMISTSLIITAIIFIASANHKVLTITAIQDATSOI----- 83  
DB 76 AALSTDAIDRVSQFDVSLDLVIREAAQKFDLLGLRLITDIASGIGEGAMALMGEEAAFI 135  
QY 84 -----KNTPTYLQDPQLGIFSNLSBITSQTTILASTTPGVKSNLQPTTVTKN 135  
DB 136 RPRSRKRGKTTT-----TSSSTSTSTTTTSTTTT----- 181  
QY 136 TTTTQPSKPTTKQKQKPNKPNDFHFEVFNPCISNNPTCWAICKRIPNKKPG 195  
DB 182 TTTTITTTTPTT----- 194  
QY 196 KTTTTPKPTKPTTKKOLKQPTTKPEVPTTKPTBEPTINTKTNTITLLTNNTGN 255  
DB 195 -----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTSTTTTSTTTT 249  
QY 256 PKLTSQMETFHSSEGNLSPQSVSTTSEHPSQSPSPNTT 296  
DB 250 PTTT-----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTTTTSTTTT 284  
  
RESULT 9  
US-09-216-393-344

US-10-321-856-341  
; Sequence 341, Application US/10321856  
; Publication No. US20030194393A1  
; GENERAL INFORMATION:  
; APPLICANT: Milhausen, Michael James  
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF  
; FILE REFERENCE: TX-1-C2-1  
; CURRENT APPLICATION NUMBER: US/10/321.856  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: 09/216.393  
; PRIOR FILING DATE: 1998-12-18  
; PRIOR APPLICATION NUMBER: 08/994.825  
; PRIOR FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 366  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 341  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Toxoplasma gondii  
US-10-321-856-341

Query Match 10.5%; Score 162.5; DB 12; Length 288;  
Best Local Similarity 23.1%; Pred. No. 1.8e-05;  
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;  
  
QY 39 AQTLSILAMISTSLIITAIIFIASANHKVLTITAIQDATSOI----- 83  
DB 76 AALSTDAIDRVSQFDVSLDLVIREAAQKFDLLGLRLITDIASGIGEGAMALMGEEAAFI 135  
QY 84 -----KNTPTYLQDPQLGIFSNLSBITSQTTILASTTPGVKSNLQPTTVTKN 135  
DB 136 RPRSRKRGKTTT-----TSSSTSTSTTTTSTTTT----- 181  
QY 136 TTTTQPSKPTTKQKQKPNKPNDFHFEVFNPCISNNPTCWAICKRIPNKKPG 195  
DB 182 TTTTITTTTPTT----- 194  
QY 196 KTTTTPKPTKPTTKKOLKQPTTKPEVPTTKPTBEPTINTKTNTITLLTNNTGN 255  
DB 195 -----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTSTTTTSTTTT 249  
QY 256 PKLTSQMETFHSSEGNLSPQSVSTTSEHPSQSPSPNTT 296  
DB 250 PTTT-----TTTPTTTTSTTTTPTTTTPTTTT-----TSSSTSTTTTSTTTT 284

RESULT 10  
US-10-321-856-344  
; Sequence 344, Application US/10321856  
; Publication No. US20030194393A1  
; GENERAL INFORMATION:  
; APPLICANT: Milhausen, Michael James  
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF  
; FILE REFERENCE: TX-1-C2-1  
; CURRENT APPLICATION NUMBER: US/10/321.856  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: 09/216.393  
; PRIOR FILING DATE: 1998-12-18  
; PRIOR APPLICATION NUMBER: 08/994.825  
; PRIOR FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 366  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 344  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Toxoplasma gondii  
US-10-321-856-344

Query Match 10.5%; Score 162.5; DB 12; Length 288;  
Best Local Similarity 23.1%; Pred. No. 1.8e-05;  
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;  
  
QY 39 AQTLSILAMISTSLIITAIIFIASANHKVLTITAIQDATSOI----- 83

Db 76 AALSTDAIDRVSQDLVSLVDVIREAAQAFDGLRLITDIASGIGEGAWALMGEEAFT 135  
Qy 84 -----KNTPTYLQDPOLGIFSGNLSEITSOITLILASTTPGVKSNLQPTTVTKKN 135  
Db 136 RPRRSKGGKTTT-----TSSSISSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 181  
Qy 136 TTTTQTPSKPTTKQKQKPNKPNNDHFVEFVPCISCSNNPTCWAICKIPNKKPG 195  
Db 182 TTTTSTTTTSTT-----194  
Qy 196 KKTTKTKPTTKTKKOLKQTKPKVEPTTKPTTEPTINTKNTTLLTNNNTGN 255  
Db 195 -TTE 249  
Qy 256 PKLTSQMETFHSSEGNLSPSQSVSTTSEHPSQSPSPNTT 296  
Db 250 PTTT-----TTTPTTTSTTTTTTTTTTTTTTTTTTTTTTTTTT 284  
  
RESULT 11  
US-10-124-557-14  
; Sequence 14, Application US/10124557  
; Publication No. US20020137894A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Katherine  
; Clark, Stephen C.  
; Jacobs, Kenneth  
; Hewick, Rodney M.  
; Gesner, Thomas G.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/124,557  
; FILING DATE: 16-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 941 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-10-124-557-14

Query Match 10.5%; Score 162; DB 14; Length 941;  
Best Local Similarity 24.6%; Pred. No. 9e-05;  
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
  
Qy 72 TTAIIQDATSOKNTTPYLVQDPOLGIFSGNLSEITSOITLILASTTPGVKSNLQPTTV 131  
Db 125 TSKVLAKTPKAEITTK-----GPAITTPKEPTTPKEPASTTP---KEPTPTTI 172  
Qy 132 K-----TKNTTTTQTPQSKPTTKQKQKPNKPNNDHFVEFVPCISCSN 178  
Db 173 KSAPPTPKPAPPTTKSAPPTPKPAPPTTKPAPPTTKPAPPTTKPAPPTTKPAP 221  
Qy 179 NPTCWAICRIPN-----KPKGKTTTKP-----TKKPTFTTK 212  
Db 222 APTTKSAPPTPKPAPPTPKKAPPTPKPAPPTTKPAPPTTKPAPPTTKPAPPTTK 281  
Qy 213 KDL-----KQPTTKPKVEPTTKP-----TEPTINTKNTTLLTNNNTGNPKLTS 260  
Db 282 EPAPTAPKPKPAPPTTKPAPPTTKPAPPTTKPAPPTTKPAPPTTKPAPPTTKPAP 341  
Qy 261 QMETFHSSEGNLSPSQSVSTTSEHPSQSPSPNTT 296  
Db 342 TTKSAPPTPKP-----PSPTTKPAPPTTKPAPPTT 373  
  
RESULT 12  
US-10-124-557-84  
; Sequence 84, Application US/10124557  
; Publication No. US20020137894A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, Katherine  
; Clark, Stephen C.  
; Jacobs, Kenneth  
; Hewick, Rodney M.  
; Gesner, Thomas G.  
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/124,557  
; FILING DATE: 16-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/643,502  
; FILING DATE: 18-JAN-1991  
; APPLICATION NUMBER: US 07/546,114  
; FILING DATE: 29-JUN-1990  
; APPLICATION NUMBER: US 07/457,196  
; FILING DATE: 29-DEC-1989  
; APPLICATION NUMBER: US 07/390,901  
; FILING DATE: 08-AUG-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cserr, Luann  
; REGISTRATION NUMBER: 31,822  
; REFERENCE/DOCKET NUMBER: GI 5190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)876-1170  
; TELEFAX: (617)876-5851  
; INFORMATION FOR SEQ ID NO: 84:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1022 amino acids  
; TYPE: amino acid

```
;
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-124-557-84

Query Match      10.5%; Score 162; DB 14; Length 1022;
Best Local Similarity 24.6%; Pred. No. 0.0001;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDASQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLOPTTV 131
Db 206 TSKVLAKPTPKAETTTK-----GPA LTTPKETPTTPKBPASTTP---KEPTPTTI 253
QY 132 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFVPCSCSN 178
Db 254 KSAPTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APT TTKBP 302
QY 179 NPTCWAICKRIPN-----KKPGKKTITTKP-----TKKPTFKITK 212
Db 303 APTTTKSAPTTTPKEPAPTTTKKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPK 362
QY 213 KDL-----KPQTTKPKVPTTKP-----TEPTINTTKNTITLLTNNTGNPKLTS 260
Db 363 EPAPTAPKAPTTTPKEPAPTTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKBPAPT 422
QY 261 QMETFHSTSGEGLNSPSQVSTTSEHPSQSPSPNNTT 296
Db 423 TTKSAPTTPKP-----PSPTTKPEPAPTTTPKEPAPTT 454

RESULT 13
US-10-124-557-74
; Sequence 74, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
;           Clark, Stephen C.
;           Jacobs, Kenneth
;           Hewick, Rodney M.
;           Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
```

```
;
;   TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 74:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1038 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match      10.5%; Score 162; DB 14; Length 1038;
Best Local Similarity 24.6%; Pred. No. 0.0001;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDASQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLOPTTV 131
Db 190 TSKVLAKPTPKAETTTK-----GPA LTTPKETPTTPKBPASTTP---KEPTPTTI 237
QY 132 K-----TKNTTTTQTPSKPTTKQKQKPNKPNNDFFHEVFNFVPCSCSN 178
Db 238 KSAPTTPKEPAPTTTKSAPTTTPKEPAPTTTPKEPAPTTTPKEP-----APT TTKBP 286
QY 179 NPTCWAICKRIPN-----KKPGKKTITTKP-----TKKPTFKITK 212
Db 287 APTTTKSAPTTTPKEPAPTTTKKAPTTTPKEPAPTTTPKEPAPTTTPKEPAPTTTPK 346
QY 213 KDL-----KPQTTKPKVPTTKP-----TEPTINTTKNTITLLTNNTGNPKLTS 260
Db 347 EPAPTAPKAPTTTPKEPAPTTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKBPAPT 406
QY 261 QMETFHSTSGEGLNSPSQVSTTSEHPSQSPSPNNTT 296
Db 407 TTKSAPTTPKP-----PSPTTKPEPAPTTTPKEPAPTT 438

RESULT 14
US-10-124-557-58
; Sequence 58, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
;           Clark, Stephen C.
;           Jacobs, Kenneth
;           Hewick, Rodney M.
;           Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:36:05 ; Search time 16.3057 Seconds  
(without alignments)  
773.267 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTLEKTDYTL.....VSTSEHPSPSPPTTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544	100.0	298	2	US-08-467-963C-8
2	1544	100.0	298	2	US-08-838-189D-8
3	1544	100.0	298	3	US-08-852-344D-8
4	1544	100.0	298	3	US-08-344-639E-8
5	1544	100.0	298	3	US-08-467-969A-8
6	1544	100.0	298	3	US-08-467-961A-8
7	1544	100.0	298	3	US-08-001-554A-8
8	901	58.4	681	6	5194595-19
9	295.5	19.1	263	5	PCT-US91-08177-13
10	222	14.4	37	3	US-08-793-792-12
11	193	12.5	32	3	US-08-793-792-8
12	188	12.2	216	3	US-08-928-361B-8
13	188	12.2	216	3	US-08-588-995A-8
14	188	12.2	1837	4	US-08-928-361B-5
15	188	12.2	1837	4	US-08-588-995A-5
16	187.5	12.1	1721	3	US-08-700-651-5
17	187.5	12.1	1721	3	US-08-928-361B-6
18	187.5	12.1	1721	3	US-08-588-995A-6
19	183.5	11.9	216	3	US-08-928-361B-27
20	174	11.3	28	3	US-08-793-792-4
21	173.5	11.2	249	3	US-08-700-651-15
22	173.5	11.2	249	3	US-08-928-361B-20
23	173.5	11.2	249	4	US-08-588-995A-20
24	169	10.9	36	3	US-09-082-279B-871
25	169	10.9	36	4	US-09-315-304B-871
26	169	10.9	36	4	US-09-834-784-871
27	164.5	10.7	2476	2	US-08-276-967-2

28 162.5 10.5 288 4 US-09-216-393B-341 Sequence 341, App  
29 162.5 10.5 288 4 US-09-216-393B-344 Sequence 344, App  
30 162 10.5 941 4 US-07-757-022B-14 Sequence 14, Appl  
31 162 10.5 1022 4 US-07-757-022B-84 Sequence 84, Appl  
32 162 10.5 1038 4 US-07-757-022B-74 Sequence 74, Appl  
33 162 10.5 1049 4 US-07-757-022B-58 Sequence 58, Appl  
34 162 10.5 1140 4 US-07-757-022B-104 Sequence 104, App  
35 162 10.5 1270 4 US-07-757-022B-44 Sequence 44, Appl  
36 162 10.5 1311 4 US-07-757-022B-42 Sequence 42, Appl  
37 162 10.5 1313 4 US-07-757-022B-142 Sequence 142, App  
38 162 10.5 1314 4 US-07-757-022B-50 Sequence 50, Appl  
39 162 10.5 1320 4 US-07-757-022B-46 Sequence 46, Appl  
40 162 10.5 1320 4 US-07-757-022B-60 Sequence 60, Appl  
41 162 10.5 1354 4 US-07-757-022B-48 Sequence 48, Appl  
42 162 10.5 1361 4 US-07-757-022B-40 Sequence 40, Appl  
43 162 10.5 1363 4 US-07-757-022B-52 Sequence 52, Appl  
44 162 10.5 1404 4 US-07-757-022B-2 Sequence 2, Appl  
45 162 10.5 1404 4 US-07-757-022B-62 Sequence 62, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-467-963C-8

; Sequence 8, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/08/467.963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 298 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-467-963C-8



Query Match 100.0%; Score 1544; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.le-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANEKVTLTIIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120  
DB 61 FIASANEKVTLTIIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNVPVCSICSNPP 180  
DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKKDLKPKQTTKPKKEVPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKKDLKPKQTTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNIITLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298  
DB 241 TNIITLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298

## RESULT 2

US-08-838-189D-8  
; Sequence 8, Application US/08838189D  
; Patent No. 5998169  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EMASISHYN, Mary E  
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,189D  
; FILING DATE: 16-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:j.b  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

## US-08-838-189D-8

Query Match 100.0%; Score 1544; DB 2; Length 298;  
Best Local Similarity 100.0%; Pred. No. 2.le-132;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANEKVTLTIIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120  
DB 61 FIASANEKVTLTIIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTOTTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNVPVCSICSNPP 180  
DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQONKPNKPNNDHFEVFNVPVCSICSNPP 180

QY 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKKDLKPKQTTKPKKEVPTTKPTTEPTINTTK 240  
DB 181 TCWAICKRIPNKKPGKTTTKPTKKPTFKTKKDLKPKQTTKPKKEVPTTKPTTEPTINTTK 240

QY 241 TNIITLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298  
DB 241 TNIITLLTNNNTGNPKLTSMETFHSTSEGNLSPSOVSTTSEHPSPQSPSPNTTRQ 298

## RESULT 3

US-08-852-344D-8  
; Sequence 8, Application US/08852344D  
; Patent No. 6017539  
; GENERAL INFORMATION:  
; APPLICANT: KLEIN, Michel H  
; APPLICANT: DU, Run-Pan  
; APPLICANT: EMASISHYN, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION  
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY  
; TITLE OF INVENTION: SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,344D  
; FILING DATE: 07-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,639  
; FILING DATE: 14-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-688 MIS:j.b  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

```
;
; TOPOLOGY: linear
; US-08-852-344D-8
;
; Query Match 100.0%; Score 1544; DB 3; Length 298;
; Best Local Similarity 100.0%; Pred. No. 2.1e-132;
; Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 MSKNKQORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
; Db 1 MSKNKQORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
;
; QY 61 FIASANKVLTALIIQDATSQIKNTPTTYLTQDPOLGIFSLSNLSITSTTTILASTTP 120
; Db 61 FIASANKVLTALIIQDATSQIKNTPTTYLTQDPOLGIFSLSNLSITSTTTILASTTP 120
;
; QY 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
; Db 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
;
; QY 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKPKVEPTTKPTTEPTINTTK 240
; Db 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKPKVEPTTKPTTEPTINTTK 240
;
; QY 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQPSPPNTTRQ 298
; Db 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQPSPPNTTRQ 298
;
; RESULT 4
; US-08-344-639E-8
; Sequence 8, Application US/08344639E
; Patent No. 6033668
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysghyn, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,639E
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-391 MIS
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELECOMMUNICATION INFORMATION:
; TELEFAX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-344-639E-8
;
; Query Match 100.0%; Score 1544; DB 3; Length 298;
; Best Local Similarity 100.0%; Pred. No. 2.1e-132;
; Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 MSKNKQORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
; Db 1 MSKNKQORTAKTEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
;
; QY 61 FIASANKVLTALIIQDATSQIKNTPTTYLTQDPOLGIFSLSNLSITSTTTILASTTP 120
; Db 61 FIASANKVLTALIIQDATSQIKNTPTTYLTQDPOLGIFSLSNLSITSTTTILASTTP 120
;
; QY 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
; Db 121 GVKSNLOPTTVTKNTTTTQTSKPTTKQKQKPNKPNNDHFVFNVPSCSNNP 180
;
; QY 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKPKVEPTTKPTTEPTINTTK 240
; Db 181 TCWAIKCRIPNKKPGKTTTKPTKPTTKKDLKPQTKPKVEPTTKPTTEPTINTTK 240
;
; QY 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQPSPPNTTRQ 298
; Db 241 TNITTTLLNTNNTGNPKLTSQMETFHSSTSEGNLSPSQSVSTTSEHPSQPSPPNTTRQ 298
;
; RESULT 5
; US-08-467-969A-8
; Sequence 8, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysghyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
```

```
;
; TELEFAX: (416) 595-1163
; TELEPHONE: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-969A-8

Query Match      100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Qy 61 FIASANHKVLTITAIQDATSOIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Db 61 FIASANHKVLTITAIQDATSOIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Qy 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Db 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Qy 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKQDKLPQTKPKVEPTTKPEEPTINTTK 240
Db 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKQDKLPQTKPKVEPTTKPEEPTINTTK 240
Qy 241 TNITTTLLTNNNTGNPKLTSQMETFHTSSEGNLSPSOVSTTSEHPQSPSPNTTRQ 298
Db 241 TNITTTLLTNNNTGNPKLTSQMETFHTSSEGNLSPSOVSTTSEHPQSPSPNTTRQ 298

RESULT 6
US-08-467-961A-8
; Sequence 8, Application US/08467961A
; Patent No. 6171783
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
```

```
;
; REFERENCE/DOCKET NUMBER: 1038-476 MIS:bh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELE: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-961A-8

Query Match      100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKDORTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Qy 61 FIASANHKVLTITAIQDATSOIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Db 61 FIASANHKVLTITAIQDATSOIKNTTPTTYLTQDPQLGISFNSLSEITSTQTTILASTTP 120
Qy 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Db 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKQKPNKPNNDFFHFEVNFVPCSCSNNP 180
Qy 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKQDKLPQTKPKVEPTTKPEEPTINTTK 240
Db 181 TCWAICRIKPNKPGKTKTKPTTKKPTTKKQDKLPQTKPKVEPTTKPEEPTINTTK 240
Qy 241 TNITTTLLTNNNTGNPKLTSQMETFHTSSEGNLSPSOVSTTSEHPQSPSPNTTRQ 298
Db 241 TNITTTLLTNNNTGNPKLTSQMETFHTSSEGNLSPSOVSTTSEHPQSPSPNTTRQ 298

RESULT 7
US-08-001-554A-8
; Sequence 8, Application US/08001554A
; Patent No. 6225091
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001,554A
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-286
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELE: 065-24567 SIMBAS
```

```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-001-554A-8

Query Match 100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQAQITLSILAMIISTLSIITAI 60
DB 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQAQITLSILAMIISTLSIITAI 60

QY 61 FIASANHKVLTITAIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTTILASTTP 120
DB 61 FIASANHKVLTITAIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTTILASTTP 120

QY 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKORQKPNKPNNDFFHFEVFNVPSCISNNP 180
DB 121 GVKSNLOPTTVTKNTTITOTQPSKPTTKORQKPNKPNNDFFHFEVFNVPSCISNNP 180

QY 181 TCWAICKRIPNPKGKTKTKPTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240
DB 181 TCWAICKRIPNPKGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 240

QY 241 TNIITTTLLTNNNTGPKLTSOMETFHSTSGEGLSPSOVSTTSEHPSQSPSPNTTRQ 298
DB 241 TNIITTTLLTNNNTGPKLTSOMETFHSTSGEGLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 8
5194595-19
; PATENT NO. 5194595
; APPLICANT: WATHEW, MICHAEL W.
; TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING
; IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY
; SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/543,780
; FILING DATE: 31-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 137,387
; FILING DATE: 23-DEC-1987
; SEQ ID NO: 19:
; LENGTH: 681
5194595-19

Query Match 58.4%; Score 901; DB 6; Length 681;
Best Local Similarity 91.0%; Pred. No. 1.6e-73;
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 96 QLIGISFSLNLSITSTTILASTTPGVKSNLOPTTVTKNTTITOTQPSKPTTKORQKP 155
DB 490 QLIGISFSPNLSITSTTILASTTPGVKSNLOPTTVTKNTTITOTQPSKPTTKORQKP 549

QY 156 PNKPNNDFFHFEVFNVPSCISNNPCTCWAICKRIPNPKGKTKTKTKTKTKTKTKOL 215
DB 550 PSKPNNDFFHFEVFNVPSCISNNPCTCWAICKRIPNPKGKTKTKTKTKTKTKKDP 609

QY 216 KPQTKPKVPTTKPTTPTNTKNTTITLLTNNNTGNPKLSQMETFHSSTSEGNLS 275
DB 610 KPQTKSKVPTTKPTTPTNTKNTTITLLTNNNTGNBELTSQMETFHSSTSEGNPS 669

QY 276 PSQVSTTSE 284
DB 670 PSQVNISSQ 678

; TITLE OF INVENTION: Antigenic peptides derived from the
```

```

RESULT 9
PCT-US91-08177-13
; SEQUENCE 13, Application PC/TUS9108177
; GENERAL INFORMATION:
; APPLICANT: Samal, Siba K
; TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08177
; FILING DATE: 19911104
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/608,937
; FILING DATE: 05-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Highest, David W
; REGISTRATION NUMBER: 30,265
; REFERENCE/DOCKET NUMBER: 20509-96711
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4854
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-08177-13

Query Match 19.1%; Score 295.5; DB 5; Length 263;
Best Local Similarity 31.7%; Pred. No. 4.1e-19;
Matches 86; Conservative 35; Mismatches 123; Indels 27; Gaps 7;

QY 1 MSKNKQRTAKTLEKTDWTLNHLFISSGLYKMLKSVQAQITLSILAMIISTLSIITAI 60
DB 1 MSNTHHPKFKTLKRAWKASKYFIVGLSCLYKFNLSLVQATATSLAMITLSLVITAI 60

QY 61 FIASANHKVLTITAIQDATSQIKNTTPTVLTQDPQLGIFSLSNLSITSTTILASTTP 120
DB 61 YISVGNKAKPTSKPTTQQTQOPQNHITPILPTEHNHKSHTSTSTUSQPNI--DTTS 118

QY 121 GVK-----SNLOPTTVTKNT--TTTQTPSKPTTKORQKPNKPNNDFFHFEVFNVPSC 174
DB 119 GTTYGHPINRTQNRKIKSQSTPLATKPLINPL----ESNPFENHQDNNSQTLPHPVCS 174

QY 175 ICSNPTCWAIC-----KIPN-----KPGKTKTKTKTKTKTKTKTKTKTK 222
DB 175 TCEGNPACSPCLQIGLERAPSRAPTITLKAPKPKTKTKTKTKTKTKTKTKTKTK 234

QY 223 KEVPTTKPTTEPTNTTKNTTITLLTNNNTT 253
DB 235 TATPQOGLISSP---EHQTNQSTTQISQHTS 262

RESULT 10
US-08-793-792-12
; SEQUENCE 12, Application US/08793792
; Patent No. 6077511
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antigenic peptides derived from the
```

;; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
;; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
;; NUMBER OF SEQUENCES: 13

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/793,792

;; FILING DATE:  
;; CLASSIFICATION: 514  
;; INFORMATION FOR SEQ ID NO: 12:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 37 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
US-08-793-792-12

Query Match 14.4%; Score 222; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 NKPNNDFHEVFNFVPCISCSNNPTCWAICKRIPNKK 193

Db 1 NKPNNDFHEVFNFVPCISCSNNPTCWAICKRIPNKK 37

## RESULT 11

US-08-793-792-8  
;; Sequence 8, Application US/08793792  
;; Patent No. 6077511

;; GENERAL INFORMATION:

;; APPLICANT:

;; TITLE OF INVENTION: Antigenic peptides derived from the

;; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
;; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
;; NUMBER OF SEQUENCES: 13

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/793,792

;; FILING DATE:

;; CLASSIFICATION: 514

;; INFORMATION FOR SEQ ID NO: 8:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 32 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

US-08-793-792-8

Query Match 12.5%; Score 193; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 4.5e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 KPNNDHEVFNFVPCISCSNNPTCWAICKRI 189

Db 1 KPNNDHEVFNFVPCISCSNNPTCWAICKRI 32

## RESULT 12

US-08-928-361B-8

;; Sequence 8, Application US/08928361B

;; Patent No. 6071518

;; GENERAL INFORMATION:

;; APPLICANT: Petersen, Carolyn

;; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

;; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

;; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

;; TITLE OF INVENTION: SPECIES INFECTIONS

;; NUMBER OF SEQUENCES: 30

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

;; STREET: 385 Sherman Avenue, Suite 6

;; CITY: Palo Alto

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94306-1840

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent in Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/928,361B

;; FILING DATE: 12-SEP-1997

;; CLASSIFICATION:

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 60/026,062

;; FILING DATE: 13-SEP-1996

;; ATTORNEY/AGENT INFORMATION:

;; NAME: VERNY, Hana

;; REGISTRATION NUMBER: 30,518

;; REFERENCE/DOCKET NUMBER: 480.76-1(HV)

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 650-324-1677

;; TELEFAX: 650-324-1678

;; INFORMATION FOR SEQ ID NO: 8:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 216 amino acids

;; TYPE: amino acid

;; STRANDEDNESS:

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

US-08-928-361B-8

Query Match 12.2%; Score 188; DB 3; Length 216;

Best Local Similarity 26.4%; Pred. No. 1.8e-09;

Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 70 TLTAIIQDATSQIKNTPTPLTQDPOLGISFNSLSEITSQTTLTILASTTPGVKSNLOPT 129

Db 12 TTTTITT 71

QY 130 TVTKNTTTTQTPSKPTTKQRONKPNKPNDFHFEVNFVPCISCSNNPTCWAICKRI 189

Db 72 TTTTITTTTTTTTTTTTTTTT-----TTTTTTTTTTTTTTTTTTTTTTTTTT 111

QY 190 PNKKPGKTTTKKPKTKKDLKPKQTKPKVPTTKPTKEPTINTTKNITLT 249

Db 112 TTKKPTTT 171

QY 250 NNTTGNPKLTOMETHTSTSEGNLSPQSVSTTSEHPSQSPSPNTT 296

Db 172 TTTTITTTTTTTTKKPTTT-----TTTTTTTTTKKPTTTTATTTT 214

## RESULT 13

US-09-588-995A-8

;; Sequence 8, Application US/09588995A

;; Patent No. 6514697

;; GENERAL INFORMATION:

;; APPLICANT: PETERSEN, CAROLYN

;; APPLICANT: BARNES, DEBRA A.

;; APPLICANT: NELSON, RICHARD C.

;; APPLICANT: GUT, JIRI

;; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND

; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM  
; FILE OF INVENTION: INFECTIONS

; FILE REFERENCE: 480.19-5

; CURRENT APPLICATION NUMBER: US/09/588,995A

; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: 08/827,171

; PRIOR FILING DATE: 1997-03-27

; PRIOR APPLICATION NUMBER: 08/928,361

; PRIOR FILING DATE: 1997-09-12

; PRIOR APPLICATION NUMBER: 08/700,651

; PRIOR FILING DATE: 1996-08-14

; PRIOR APPLICATION NUMBER: 08/415,751

; PRIOR FILING DATE: 1995-04-03

; NUMBER OF SEQ ID NOS: 115

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 216

; TYPE: PRT

; ORGANISM: Cryptosporidium parvum

US-09-588-995A-8

Query Match 12.2%; Score 188; DB 4; Length 216;

Best Local Similarity 26.4%; Pred. No. 1.8e-09;

Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;

QY 70 TLTATIQDQTSQIKNTPTYLQDPQLGIFSNLSEITSTQTTILASTTPGVKSNLOPT 129

DB 12 TTTTITT 71

QY 130 TVTKNTNTTQPSKPTTKQKQKPNKPNDFHFEVFNVPSCSNPTCWAICKRI 189

DB 72 TTTTITT 111

QY 190 PNKKPGKTK 249

DB 112 TTKKPTTT 171

QY 250 NNTGNPKLTQSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTT 296

DB 172 TTTTITTTTTTATTTTKKPTTT-----TTTTTTTTTTTKKPTTTTATTTT 214

## RESULT 14

US-08-928-361B-5

Sequence 5, Application US/08928361B

Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn

TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

TITLE OF INVENTION: SPECIES INFECTIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESSES:

ADDRESSES: PETERS, VERNY, JONES & BIKSA

STREET: 385 Sherman Avenue, Suite 6

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,361B

FILING DATE: 12-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,062

FILING DATE: 13-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Verny, Hana

REGISTRATION NUMBER: 30,518

REFERENCE/DOCKET NUMBER: 480.76-1 (HV)

TELEPHONE: 650-324-1677

TELEFAX: 650-324-1678

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1837 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-928-361B-5

Query Match

Best Local Similarity 22.7%; Pred. No. 3.5e-08;

Matches 70; Conservative 32; Mismatches 161; Indels 46; Gaps 5;

QY 9 TaktLEKtWdTLNHLFISSGLYKLNKSVAQITISILAMIIST-----52

DB 239 TUKTIPNTY-----AGVRSNETKTTFPSANTNELLVDPKINAPCNSSEPEQVQ 288

QY 53 -----SLIITAIIFIASANNKVTLTATIQDQTSQIKNTPTVLTQDPQLGIFSNLSEI 107

DB 289 IFDMGSKVPIPTKCVGVKHTTTTTTTTTTTTTTTTTTTTTTTTTTTT-----TTT 339

QY 108 TSCQTTILASTTPGVKSNLOPTTVTKNTTTTOTQPSKPTTKQKQKPNKPNNDPHEV 167

DB 340 TTTTITT-----TTT 397

QY 168 FNFVPCSCSNPTCWAICKRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 227

DB 398 TTTTITT-----TTT 457

QY 228 TKPTBEPINTTKNITTTLLNNNTGNPKLTQSMETFHSTSEGNLSPSOVSTSEHPS 287

DB 458 TTTTITT-----TTT 513

QY 288 QPSSPNTT 296

DB 514 TTTTATTTT 522

## RESULT 15

US-09-588-995A-5

Sequence 5, Application US/09588995A

Patent No. 6514697

GENERAL INFORMATION:

APPLICANT: PETERSEN, CAROLYN

APPLICANT: BARNES, DEBRA A.

APPLICANT: NELSON, RICHARD C.

APPLICANT: GUT, JIRI

TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND

TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM

TITLE OF INVENTION: INFECTIONS

FILE REFERENCE: 480.19-5

CURRENT APPLICATION NUMBER: US/09/588,995A

CURRENT FILING DATE: 2000-06-06

PRIOR APPLICATION NUMBER: 08/827,171

PRIOR FILING DATE: 1997-03-27

PRIOR APPLICATION NUMBER: 08/928,361

PRIOR FILING DATE: 1997-09-12

PRIOR APPLICATION NUMBER: 08/700,651

PRIOR FILING DATE: 1996-08-14

PRIOR APPLICATION NUMBER: 08/415,751

PRIOR FILING DATE: 1995-04-03

NUMBER OF SEQ ID NOS: 115

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 1837

TYPE: PRT

ORGANISM: Cryptosporidium parvum

Search completed: October 29, 2003, 17:42:22  
Job time : 17.3057 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:11:05 ; Search time 68.034 Seconds  
(without alignments)  
695.249 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTLEKTDL.....VSTSEHPSQSPSPNTTRQ 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1544	100.0	298	14 AAR39286	Respiratory syncyt
2	1544	100.0	298	20 AAW96313	Membrane bound G p
3	1418	91.8	298	8 AAR70845	Sequence of human
4	1418	91.8	298	13 AAR25302	HSRV glycoprotein
5	1418	91.8	298	19 AAW47605	HSRV glycoprotein
6	1418	91.8	298	23 AAU74676	Respiratory syncyt
7	1229	79.6	232	20 AAW96314	Secreted G protein
8	901	58.4	681	10 AAR90441	Chimeric human res
9	771	49.9	299	22 AAB68336	Amino acid sequenc

10	580	37.6	361	24 AAB97862	RSV G protein pept
11	572.5	37.1	452	22 AAB67771	Amino acid sequenc
12	564.5	36.6	291	24 AAB97861	RSV G protein pept
13	564.5	36.6	548	24 AAB97863	RSV G protein pept
14	562	36.4	349	17 AAB95660	Streptococcal prot
15	562	36.4	349	22 AAB68028	Amino acid sequenc
16	558	36.1	101	16 AAB88253	RSV subgroup A wil
17	558	36.1	101	17 AAB95610	RSV subgp. A prote
18	558	36.1	101	17 AAB95616	RSV sub-group A wi
19	558	36.1	101	17 AAB97050	Respiratory Syncyt
20	558	36.1	101	20 AAY44078	RSV G protein anti
21	558	36.1	101	21 AAB18805	A G2Na peptide der
22	558	36.1	101	22 AAG67741	Amino acid sequenc
23	558	36.1	101	22 AAB84123	Amino acid sequenc
24	558	36.1	101	22 AAB68016	Amino acid sequenc
25	558	36.1	101	22 AAB67775	Amino acid sequenc
26	558	36.1	101	23 AAO22577	Human G protein, G
27	552.5	35.8	356	17 AAB95661	Immunogenic carrie
28	538	34.8	101	16 AAB88255	RSV subgroup A mod
29	538	34.8	101	17 AAB95612	RSV subgp. A prote
30	538	34.8	101	17 AAB95618	RSV sub-group A mu
31	538	34.8	101	17 AAB97052	Respiratory Syncyt
32	538	34.8	101	20 AAY44080	RSV G protein anti
33	538	34.8	101	20 AAW97311	Peptide which indu
34	538	34.8	101	22 AAB84125	Amino acid sequenc
35	538	34.8	101	23 AAO22581	Human G protein, G
36	538	34.8	101	23 AAO22586	Human G protein, G
37	534	34.6	101	20 AAW97310	Peptide which indu
38	534	34.6	101	23 AAO22585	Human G protein, G
39	514	33.3	101	20 AAW97312	Peptide which indu
40	514	33.3	101	23 AAO22587	Human G protein, G
41	506	32.8	101	17 AAB95614	RSV subgp. A prote
42	506	32.8	101	17 AAB97063	Respiratory Syncyt
43	506	32.8	101	20 AAY44090	RSV G protein anti
44	506	32.8	101	22 AAB84135	Amino acid sequenc
45	506	32.8	101	23 AAO22584	Human G protein, G

#### ALIGNMENTS

#### RESULT 1

AAR39286  
ID AAR39286 standard; Protein; 298 AA.

XX AAR39286;

XX 25-MAR-2003 (updated)

DT 13-JAN-1994 (first entry)

XX Respiratory syncytial virus (RSV) G protein.

XX PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.

OS Respiratory syncytial virus.

XX WO9314207-A1.

PD 22-JUL-1993.

PF 05-JAN-1993; 93WO-CA00001.

PR 06-JAN-1992; 92GB-0000117.

XX (CONN-) CONNAUGHT LAB LTD.

XX Ewasysbyn ME, Klein MH;

XX WPI; 1993-243222/30.

XX N-PSDB; AAQ45686.

PT Multimeric hybrid genes and their chimeric proteins - are  
PT vaccines against multiple pathogenic infections e.g.



PT para-influenza virus and respiratory syncytial virus  
 XX Claim 11; Figure 7A-7D; 80pp; English.

XX A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 298 AA;  
 SQ Query Match 100.0%; Score 1544; DB 14; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-113;  
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 DB 1 MSKNKQORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 QY 61 FIASANKHVTLTATIQDATSQIKNTTPTLTQDPQLGIFSFLSILSITSTTTILASTTP 120  
 DB 61 FIASANKHVTLTATIQDATSQIKNTTPTLTQDPQLGIFSFLSILSITSTTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFNFPVCSICSNNP 180  
 DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFNFPVCSICSNNP 180  
 QY 181 TCWAICKRIPNKKPGKNTTKKPTTKTKKDLKPKQTKKEVPTTKPTBPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKNTTKKPTTKTKKDLKPKQTKKEVPTTKPTBPTINTTK 240  
 QY 241 TMTITLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 DB 241 TMTITLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 2

AAW96313  
 ID AAW96313 standard; Protein; 298 AA.  
 XX  
 AC AAW96313;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein of respiratory syncytial virus.  
 DE  
 DE G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 XW tissue plasminogen activator.  
 XX

OS Respiratory syncytial virus.

XX WO9904010-A1.

XX 28-JAN-1999.

XX 16-JUL-1998; 98WO-CA00697.

XX 18-JUL-1997; 97US-0896442.

XX (CONN-) CONNAUGHT LAB LTD.

XX Klein MH, Li X, Sambhara S;

XX WPI; 1999-132254/11.

DR N-PSDB; AAX08421.

XX Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector

PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX Claim 4; Fig 2; 67pp; English.

XX The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.

SQ Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 20; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-113;  
 Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 DB 1 MSKNKQORTAKTLEKWTDLNHLFISSGLYKLNKLSVAQITLSILAMIIISTLIITAI 60  
 QY 61 FIASANKHVTLTATIQDATSQIKNTTPTLTQDPQLGIFSFLSILSITSTTTILASTTP 120  
 DB 61 FIASANKHVTLTATIQDATSQIKNTTPTLTQDPQLGIFSFLSILSITSTTTILASTTP 120  
 QY 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFNFPVCSICSNNP 180  
 DB 121 GVKSNLOPTTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEVFNFPVCSICSNNP 180  
 QY 181 TCWAICKRIPNKKPGKNTTKKPTTKTKKDLKPKQTKKEVPTTKPTBPTINTTK 240  
 DB 181 TCWAICKRIPNKKPGKNTTKKPTTKTKKDLKPKQTKKEVPTTKPTBPTINTTK 240  
 QY 241 TMTITLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 DB 241 TMTITLLTNNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 3

AAP70845  
 ID AAP70845 standard; protein; 298 AA.  
 XX  
 AC AAP70845;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-APR-1991 (first entry)  
 XX  
 DE Sequence of human respiratory syncytial virus (HRSV) A2 strain  
 DE G protein.  
 XX  
 KW Vaccine.  
 XX  
 XX Human respiratory syncytial virus (HRSV).  
 XX  
 XX WO8704185-A.  
 XX  
 XX 16-JUL-1987.  
 XX  
 XX 23-DEC-1986; 86WO-US02756.  
 XX  
 XX 14-JAN-1986; 86US-0818740.  
 XX  
 XX (UNIC-) UNIV NORTH CAROLINA.  
 XX  
 XX (WERT/) WERTZ G W.

```

DR WPI; 1987-206300/29.
DR N-PSDB; AAN70784.
XX Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
XX Disclosure; Chart 13; 57pp; English.
XX A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 298 AA;
SQ Query Match 91.8%; Score 1418; DB 8; Length 298;
Best Local Similarity 93.3%; Pred. No. 2e-103;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKQORTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVTLTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120
Db 61 FIASANKHVTPTTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPVCSICSNNP 180
Db 121 GVKSTLQSTTVTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPVCSICSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPKPTTKDLPQTTKPKVPTTKPTTEPTINTTK 240
Db 181 TCWAICKRIPNKKPGKTTTKPKPTTKDLPQTTKPKVPTTKPTTEPTINTTK 240

QY 241 TNIITTTLLNTGNPKLTSQMETFHTSSEGNLSFQSVSTTSEHSPQSSPPNTRQ 298
Db 241 TNIITTTLLNTGNPELTQMETFHTSSEGNLSPSQSVSTTSEYPSQSPSPNTRQ 298

RESULT 4
AAR25302
ID AAR25302 standard; Protein; 298 AA.
XX AC AAR25302;
XX AC AAR25302;
DT 25-MAR-2003 (updated)
DT 03-MAR-1993 (first entry)
XX DE HRSV glycoprotein G (gpG).
XX KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KW major capsid protein; N.
XX OS Human respiratory syncytial virus strain A2.
XX PN US5149650-A.
XX PD 22-SEP-1992.
XX PF 13-JUL-1988; 88US-0218737.
XX PR 14-JAN-1986; 86US-0818740.
PR 13-JUL-1988; 88US-0218737.
XX (UYN-) UNIV NORTH CAROLINA.
XX PI Collins PL, Wertz GW;
XX WPI; 1992-340247/41.

DR WPI; 1987-206300/29.
DR N-PSDB; AAN70784.
XX Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
XX Disclosure; Page 18; 21pp; English.
XX The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX Sequence 298 AA;
SQ Query Match 91.8%; Score 1418; DB 13; Length 298;
Best Local Similarity 93.3%; Pred. No. 2e-103;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNKQORTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60
Db 1 MSKNKQORTAKTLEKTWDTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60

QY 61 FIASANKHVTLTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120
Db 61 FIASANKHVTPTTALIQDATSQIKNTPTTYLTQDPOLGISFNSLSEITSTTILASTTP 120

QY 121 GVKSNLQPTTVTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPVCSICSNNP 180
Db 121 GVKSTLQSTTVTKNTTTTQPSKPTTKORQNKPNKPNNDHFEVFNVPVCSICSNNP 180

QY 181 TCWAICKRIPNKKPGKTTTKPKPTTKDLPQTTKPKVPTTKPTTEPTINTTK 240
Db 181 TCWAICKRIPNKKPGKTTTKPKPTTKDLPQTTKPKVPTTKPTTEPTINTTK 240

QY 241 TNIITTTLLNTGNPKLTSQMETFHTSSEGNLSFQSVSTTSEHSPQSSPPNTRQ 298
Db 241 TNIITTTLLNTGNPELTQMETFHTSSEGNLSPSQSVSTTSEYPSQSPSPNTRQ 298

RESULT 5
AAW47605
ID AAW47605 standard; Protein; 298 AA.
XX AC AAW47605;
XX AC AAW47605;
DT 11-JUN-1998 (first entry)
XX DE HRSV glycoprotein G.
XX KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.
XX OS Human respiratory syncytial virus.
XX PN US5716823-A.
XX PD 10-FEB-1998.
XX PF 12-MAY-1997; 97US-0854783.
XX PR 13-JUL-1988; 88US-0218737.
PR 14-JAN-1986; 86US-0818740.
PR 23-DEC-1986; 86WO-US02756.
PR 11-JUN-1992; 92US-0897171.
PR 12-MAY-1997; 97US-0854783.
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX PI Collins PL, Wertz GW;
XX WPI; 1998-144802/13.

```

DR N-PSDB; AAV18736.  
 XX Production of human respiratory syncytial virus glyco-protein F or G  
 PT - by culturing eukaryotic host cells transfected with corresponding  
 PT DNA  
 XX  
 XX Example 1; Columns 27-28; 17pp; English.  
 PS  
 XX The present sequence was used in the development of a novel method  
 CC for the production of human respiratory syncytial virus (HRSV)  
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises  
 CC culturing eukaryotic host cells transfected with an isolated DNA  
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare  
 CC vaccines against HRSV.  
 XX  
 XX Sequence 298 AA;  
 SQ  
 Query Match 91.8%; Score 1418; DB 19; Length 298;  
 Best Local Similarity 93.3%; Pred. No. 2e-103;  
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 DB 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 QY 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQNPQLGSPNSPSEITTSQITTLASTTP 120  
 DB 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQNPQLGSPNSPSEITTSQITTLASTTP 120  
 QY 121 GVKSNLOPTTVKTKNTTTTQPSKPTTKORONKPNKPNNDHFEVFNVPSCISNNP 180  
 DB 121 GVKSTLQSTTVKTKNTTTTQPSKPTTKORONKPNKPNNDHFEVFNVPSCISNNP 180  
 QY 181 TCWAICRIPNKKPKGKTKTKPTTKPTTKKPKQTTKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
 DB 181 TCWAICRIPNKKPKGKTKTKPTTKPTTKKPKQTTKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
 QY 241 TNIITLLTNNTGNPKLTQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTRQ 298  
 DB 241 TNIITLLTNNTGNPKLTQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTRQ 298  
 RESULT 6  
 AAU74676  
 ID AAU74676 standard; Protein; 298 AA.  
 AC AAU74676;  
 DT 09-APR-2002 (first entry)  
 DE Respiratory syncytial virus G protein.  
 KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;  
 KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
 KW antiviral chemotherapeutic compound; humoral response;  
 KW cellular immune response; hPIV; paediatric respiratory disease;  
 KW globin gene transfer; sickle cell disease; beta-thalassaemia;  
 KW human immunodeficiency virus infection; HIV.  
 XX  
 OS Human respiratory syncytial virus.  
 XX  
 PN WO200192548-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 22-MAY-2001; 2001WO-US16610.  
 XX  
 PR 01-JUN-2000; 2000US-208701P.  
 XX  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Portner A, Takimoto T;  
 XX

DR WPI; 2002-130534/17.  
 XX N-PSDB; AAS21045.  
 PT Recombinant Sendai virus useful in vaccines to protect infection by  
 PT paramyxoviruses, comprises exogenous nucleic acid encoding  
 PT paramyxovirus protein or its antigenic fragment  
 XX  
 XX Disclosure; Page 48; 57pp; English.  
 PS  
 XX The invention relates to a recombinant Sendai virus comprising an  
 CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 CC antigenic fragment. The virus may be administered in combination  
 CC with an antiviral chemotherapeutic compound. Two or more viruses  
 CC expressing different PMV proteins may be co-administered. Compositions  
 CC comprising the virus are useful for eliciting a humoral and/or  
 CC cellular immune response to a PMV in a mammal, particularly a human.  
 CC Further a recombinant Sendai virus comprising an exogenous nucleic acid  
 CC encoding a second PMV protein is also administered and priming and/or  
 CC boosting humoral or cellular immune response comprises administering  
 CC one or more of a recombinant or isolated PMV protein or its antigenic  
 CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
 CC vector encoding a PMV protein. The recombinant virus is useful as an  
 CC effective vaccine against hPIV or RSV (the major causes of paediatric  
 CC respiratory disease) and also to express any gene of  
 CC interest in target cells, providing a positive medical impact on  
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 CC into stem cells effects a cure for sickle cell disease or beta-  
 CC thalassaemia. The recombinant virus may also prove effective in  
 CC conferring immunity to human immunodeficiency virus (HIV) infection.  
 CC The Sendai virus replicates at level that is high enough to  
 CC induce sufficient immunity, but does not cause any harm to human  
 CC recipient. The present sequence represents a respiratory syncytial  
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 CC suitable for expression by the recombinant virus of the invention.  
 XX  
 XX Sequence 298 AA;  
 SQ  
 Query Match 91.8%; Score 1418; DB 23; Length 298;  
 Best Local Similarity 93.3%; Pred. No. 2e-103;  
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 DB 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMIISTSLIITAI 60  
 QY 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQNPQLGSPNSPSEITTSQITTLASTTP 120  
 DB 61 FIASANHKVTITTAIIQDATSQIKNTTPTVLTQNPQLGSPNSPSEITTSQITTLASTTP 120  
 QY 121 GVKSNLOPTTVKTKNTTTTQPSKPTTKORONKPNKPNNDHFEVFNVPSCISNNP 180  
 DB 121 GVKSTLQSTTVKTKNTTTTQPSKPTTKORONKPNKPNNDHFEVFNVPSCISNNP 180  
 QY 181 TCWAICRIPNKKPKGKTKTKPTTKPTTKKPKQTTKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
 DB 181 TCWAICRIPNKKPKGKTKTKPTTKPTTKKPKQTTKQDKLPQTTKPKVEPTTKTEPTINTTK 240  
 QY 241 TNIITLLTNNTGNPKLTQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTRQ 298  
 DB 241 TNIITLLTNNTGNPKLTQMETFHSSTSEGNLSPSQVSTTSEHPSPQSPSPNTTRQ 298  
 RESULT 7  
 AAU96314  
 ID AAU96314 standard; Protein; 232 AA.  
 XX  
 AC AAU96314;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Secreted G protein of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO9904010-A1.  
 XX  
 PD 28-JAN-1999.  
 XX  
 XX 16-JUL-1998; 98WO-CA00697.  
 PF  
 XX 18-JUL-1997; 97US-0896442.  
 PR  
 XX (CONN-) CONNAUGHT LAB LTD.  
 PA  
 XX Klein MH, Li X, Sambhara S;  
 PI WPI; 1999-132254/11.  
 XX N-PSDB; AAX08422.  
 DR  
 XX Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 9; Fig 3; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods. This truncated G  
 CC protein is secreted since it lacks a transmembrane domain.  
 XX  
 SQ Sequence 232 AA;  
 Query Match 79.6%; Score 1229; DB 20; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 1e-88;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 67 HKVTLTALIQDQTSQIKNTPTLYLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
 DB 1 HKVTLTALIQDQTSQIKNTPTLYLTQDPQLGIFSLSNLSITSTQTTILASTTPGVKSNL 60  
 QY 127 QPTTVTKNTTQTQPSKPTTKORONKPNKNDPFHEVFNVPSCSNNPTCWAIC 186  
 DB 61 QPTTVTKNTTQTQPSKPTTKORONKPNKNDPFHEVFNVPSCSNNPTCWAIC 120  
 QY 187 KRIPNKKPGKTKTKPTTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTKNITTT 246  
 DB 121 KRIPNKKPGKTKTKPTTKTKKOLKPKQTTKPKVPTTKPTTEPTINTTKNITTT 180  
 QY 247 LITNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 DB 181 LITNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 RESULT 8  
 AAP90441  
 ID AAP90441 standard; protein; 681 AA.  
 XX  
 AC AAP90441;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX

DE Chimeric human respiratory syncytial virus glycoproteins F and G.  
 XX  
 KW Chimeric polypeptide; human respiratory syncytial virus;  
 KW protein F; protein G; vaccine.  
 XX  
 OS Human respiratory syncytial virus.  
 XX  
 PN WO8905823-A.  
 XX  
 PD 29-JUN-1989.  
 XX  
 XX 31-OCT-1988; 88WO-US03784.  
 PF  
 XX 23-DEC-1987; 87US-0137387.  
 PR  
 XX (UPJO ) UPJOHN CO.  
 PA  
 XX Mathen M;  
 PI WPI; 1989-206593/28.  
 XX  
 DR Chimeric human respiratory syncytial virus polypeptides(s)  
 PT - contg. immunogenic fragments from HRSV glycoproteins  
 PT F and G, for vaccine prodn.  
 XX  
 PS Claim 3; page 47-48; 50pp; English.  
 XX  
 CC Chimeric polypeptide contg. a signal sequence and one or more  
 CC immunogenic fragments from both human respiratory syncytial virus  
 CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg  
 CC E. coli, Chinese hamster ovary cells, murine C127 cells and  
 CC S. frugiperda.  
 XX  
 SQ Sequence 681 AA;  
 Query Match 58.4%; Score 901; DB 10; Length 681;  
 Best Local Similarity 91.0%; Pred. No. 2.1e-62;  
 Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
 QY 96 QLGISFSNLSITSTQTTILASTTPGVKSNLQPTTVTKNTTQTQPSKPTTKORONKP 155  
 DB 490 QLGISFSNLSITSTQTTILASTTPGVKSNLQPTTVTKNTTQTQPSKPTTKORONKP 549  
 QY 156 PNKPNNDHFEVFNVPSCSNNPTCWAICKRIPNKPGKTKTKPTTKPKTKKOL 215  
 DB 550 PSKPNNDHFEVFNVPSCSNNPTCWAICKRIPNKPGKTKTKPTTKPKTKKOP 609  
 QY 216 KPQTTKPEVPTTKPTTEPTINTTKNITTTLLTNTTGNPKLTSQMETFHTSSSEGNLS 275  
 DB 610 KPQTTKSEVPTTKPTTEPTINTTKNITTTLLTNTTGNPKLTSQMETFHTSSSEGNLS 669  
 QY 276 PSQVSTTSE 284  
 DB 670 PSQVSNISQ 678  
 RESULT 9  
 AAB68336  
 ID AAB68336 standard; protein; 299 AA.  
 XX  
 AC AAB68336;  
 XX  
 DT 09-JUL-2001 (first entry)  
 XX  
 DE Amino acid sequence of RSV G-protein.  
 XX  
 KW Respiratory syncytial virus; RSV; G-protein; annexin II; L-selectin;  
 KW RSV infection.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO200129054-A2.  
 XX

PD 26-APR-2001.  
 XX 23-OCT-2000; 2000WO-GB04084.  
 XX 21-OCT-1999; 99GB-0024990.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Malhotra R, Bird M;  
 XX WPI; 2001-316238/33.  
 XX Treatment of respiratory syncytial virus infection (RSV) using Annexin  
 XX II or L-selectin derivatives -  
 XX Disclosure; Page 71-73; 74pp; English.  
 XX The specification describes a product which antagonizes binding of  
 XX respiratory syncytial virus (RSV) G-protein to annexin II or  
 XX L-selectin, or causes a decrease in cell surface levels of annexin  
 XX II or L-selectin. Such products are used in the manufacture of a  
 XX medicament for use in prevention or treatment of RSV infection.  
 XX The present sequence represents a RSV G-protein.  
 XX Sequence 299 AA;  
 SQ Query Match 49.9%; Score 771; DB 22; Length 299;  
 Best Local Similarity 54.7%; Pred. No. 1.3e-52;  
 Matches 162; Conservative 36; Mismatches 96; Indels 2; Gaps 2;  
 QY 1 MSKKNQRTATLEKTDWTLNHLFISSGLYKLNKVAQITLSTLAMIISTSLIITAI 60  
 DB 1 MSKKNQRTATLEKTDWTLNHLFISSGLYKLNKVAQITLSTLAMIISTSLIITAI 60  
 QY 61 FIASANKVLTITLAIQDATSQIKNTTPTVLTQDPQLGISFNLSEITSTTTILASTTP 120  
 DB 61 FIASANKVLTITVTVQIKHTEKNITTYLTQVPPERVSSKQPTTTSPIHTNSATSP 120  
 QY 121 GVKNLQPTTKKNTTQTPSPKPTTKQKQKPNKPNNDHFVFNVPFCISCSNNP 180  
 DB 121 NTKSETHHTAQTKGRITTSQTNNKPKPRLKPNPKPKDDYHFEVFNVPFCISGNNQ 180  
 QY 181 TCWAICKPIPKKCKTTPKTKPKTKT-KKDLKPTTKPKKVPPTTKTEPTINTT 239  
 DB 181 LCKSICKTIPKPKPKTKTIPKTKPKTKTKTKPKTKTKTKTKTKTKTKTKTKTK 240  
 QY 240 KNTITTTLLTNNTPGNPKLTQSMETFTSSEGNLSPSQVSTTSHSPQSPPT 295  
 DB 241 ERDTSQSTVLDTTLEHTIQOQSLHSTTPENTPNSQTTPASE-PSTSNSTQNT 295  
 RESULT 10  
 ABP97862  
 ID ABP97862 standard; protein; 361 AA.  
 AC ABP97862;  
 XX 03-JUN-2003 (first entry)  
 DT RSV G protein peptide G2Na and diphtheria anatoxin peptide DTb.  
 DE Diphtheria anatoxin; immune response; antigen; carrier; G protein;  
 KW hypersensitivity response; vaccine; infection; RSV; cancer.  
 XX Synthetic.  
 OS Respiratory syncytial virus.  
 OS Corynebacterium diphtheriae.  
 XX FR2827606-A1.  
 XX 24-JAN-2003.  
 PD 20-JUL-2001; 2001FR-0009733.  
 PF

XX 20-JUL-2001; 2001FR-0009733.  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX Corvaia N, Nguyen NT, Beck A;  
 XX WPI; 2003-241969/25.  
 XX New peptide derived from diphtheria anatoxin, useful as carrier in  
 XX vaccines, lacks at least one Cys residue, also related nucleic acids -  
 XX Disclosure; Page 27-28; 42pp; French.  
 XX The present sequence represents a fusion protein comprising a peptide  
 XX derived from the G protein of Respiratory syncytial virus (RSV) linked  
 XX to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin  
 XX peptide generates an immune response against any antigen coupled to it,  
 XX and has a negative, immediate hypersensitivity response. Diphtheria  
 XX anatoxin derived peptides, when modified to lack at least one cysteine  
 XX residue, are useful as carrier peptides. Deletion of Cys residues in  
 XX anatoxin peptides reduces formation of unwanted disulfide bridges. The  
 XX peptides are used as a carrier for vaccines, particularly those for  
 XX prevention or treatment of viral, bacterial, parasitic or fungal  
 XX infections, or cancers and to generate, or increase, an immune response  
 XX against infectious agents or tumour cells.  
 XX Sequence 361 AA;  
 SQ Query Match 37.6%; Score 580; DB 24; Length 361;  
 Best Local Similarity 70.8%; Pred. No. 1.6e-37;  
 Matches 114; Conservative 5; Mismatches 20; Indels 22; Gaps 3;  
 QY 130 TVTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVFNVPFCISCSNNPTCWAICKRI 189  
 DB 4 TVTKNTTTTQTPSKPTTKQKQKPNKPNNDHFVFNVPFCISCSNNPTCWAICKRI 63  
 QY 190 PNKPKGKTTTKTKPKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 241  
 DB 64 PNKPKGKTTTKTKPKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 120  
 QY 242 NITF-----TLLTNTTGNPKLTQSMETFTSSE 271  
 DB 121 KIESLKEHGPIKMKMSSEPNKTVSEKAKQYLEEFHOTALE 161  
 RESULT 11  
 AAB67771  
 ID AAB67771 standard; Protein; 452 AA.  
 AC AAB67771;  
 XX 11-JUN-2001 (first entry)  
 DT Amino acid sequence of a fusion protein of P40 and RSV antigen.  
 DE Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;  
 KW RSV; RSV infection; lung; respiratory tract; vaccine.  
 XX Synthetic.  
 OS Klebsiella pneumoniae.  
 OS Respiratory syncytial virus.  
 XX WO200121203-A1.  
 XX 29-MAR-2001.  
 XX 22-SEP-2000; 2000WO-FR02626.  
 XX 23-SEP-1999; 99FR-0011888.  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:18:11 ; Search time 1944.71 Seconds  
(without alignments)  
8935.871 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715

Sequence: 1 cacaaagtcacactaacac.....gtagttattataaaaaaaaaa 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estbta:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hcc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	120	16.8	895	29	CNS0071A
C 2	113	15.8	712	13	AL066286 Drosophila
C 3	111.4	15.6	1180	13	BX416727 BX416727
C 4	104.6	14.6	946	29	BX436369 BX436369
					AG127412 Pan trogl

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 5	103.2	14.4	922	29	CNS0073W
C 6	101.8	14.2	919	29	CNS04ENY
C 7	100	14.0	1201	9	AL581589
C 8	99.8	14.0	1201	9	AL531977
C 9	99	13.8	1101	29	CNS00EXE
C 10	96.2	13.5	753	13	EX443342
C 11	96.2	13.5	1100	13	EX385076
C 12	95.8	13.4	1097	29	CNS01224
C 13	95	13.3	791	29	CNS009KS
C 14	95	13.3	1101	29	CNS00LOO
C 15	94.6	13.2	899	13	EX436853
C 16	94.4	13.2	1216	29	AG135357
C 17	94.2	13.2	815	28	B14686
C 18	94.2	13.2	1200	9	AL548181
C 19	94	13.1	1000	13	BX418086
C 20	94	13.1	1151	14	CD501049
C 21	93.8	13.1	1008	29	AG137085
C 22	93.6	13.1	884	29	CNS006UO
C 23	93.6	13.1	969	28	AQ743309
C 24	93.4	13.1	1172	29	BZ696157
C 25	93.2	13.0	1141	28	AQ743305
C 26	92.8	13.0	981	29	AG127518
C 27	92.8	13.0	1183	29	BZ696114
C 28	92.6	13.0	923	13	BX410248
C 29	92.6	13.0	1183	29	AG136828
C 30	92.4	12.9	859	29	AG128925
C 31	92.4	12.9	880	29	AG139490
C 32	92	12.9	1137	28	AQ743326
C 33	92	12.9	1225	29	CNS0166K
C 34	91.4	12.8	700	29	AG127423
C 35	91.4	12.8	885	29	CNS03LTM
C 36	91	12.7	1099	28	AQ743360
C 37	90.8	12.7	885	13	BX425603
C 38	90.8	12.7	1081	29	AG135328
C 39	90.6	12.7	1491	10	BE882936
C 40	90.2	12.6	866	29	AG126308
C 41	90	12.6	953	29	CNS008PF
C 42	90	12.6	1039	13	BX378025
C 43	90	12.6	1101	29	CNS017FC
C 44	89.8	12.6	896	29	AG141027
C 45	89.8	12.6	1201	13	BX399683

## ALIGNMENTS

RESULT 1  
CNS0071A/c 895 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
DEFINITION BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL066286  
VERSION AL066286.1 GI:4945153  
KEYWORDS GSS  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 895)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammos in Pieter de Jong's laboratory in the Department of



AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster S16.f For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOCAP001BG06QP1&cluster=S16.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOCAP001BG06QP1&cluster=S16.f). Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOCAP001BG06QP1.

FEATURES Location/Qualifiers  
source 1..1180  
organism="Homo sapiens"  
mol\_type="mRNA"  
db\_xref="taxon:9606"  
clone="CSOCAP001YM12"  
tissue type="THYMUS"  
clone\_lib="Homo sapiens THYMUS"  
note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector."  
Library was not normalized."

BASE COUNT 127 a 154 c 191 g 420 t 288 others  
ORIGIN

Query Match 15.6%; Score 111.4; DB 13; Length 1180;  
Best Local Similarity 25.8%; Pred. No. 2.6e-08;  
Matches 144; Conservative 194; Mismatches 219; Indels 1; Gaps 1;

QY 120 AATTGATCATCAAAACACCACCTACTAGTTCTTAACAACCCAGGAGTCGAAGTCAAACCT 179  
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
1152 RAAAATAAAGAANAAAAAAARFAAMAASASAAMVAAMAAAAAAAVAAAGAAVMMAAAA 1093  
QY 180 GCACACCACACAGCTCAAGACTAAAAACACCAACCAACCCCAACCAACCCAGCAGCC 239  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
1092 AAAAATAAAAAAAAAAGMAAASMSMMMMMAAASAAAAAASAAAVMAAAAAAAMA 1033  
QY 240 CACTACAAAACACGCCAAAACAAACACCAACCAACCAACCAATATGATTTTCACCTCGA 299  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
1032 AAGAAMAWAAAAANWMAAANWMAAANWMAAANWMAAANWMAAANWMAAANWMAA 973  
QY 300 AGTGTTTAAC TTGTGCCTGGCAGCATATGCGAGCAACAATCCAACCTGCTGGGCTATCTG 359  
Db ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
972 AAAAAAATAAAAKHAANAATAAAATAAYMAAMAFAAMAATAAAWMAAHNAHMAAANH 913  
QY 360 CAAGAAGATACCAACAAAAAACAGAGAAAGAAAAACCAACCAACCAAGCTCAAAAAACC 419  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
912 AMWAAMANMANMABACMCMAATPMHHMMMACNCHMMHMNMCMNMNMNMCMNWMA 853  
QY 420 AACCTTCAAGCAACCAAAAAAGATCTCAAAACCTCAACACCACTAAACCAAGCAAGTAGC 479  
Db |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
852 AAAAAAATAAAKHAANAATAAAATAAYMAAMAFAAMAATAAAWMAAHNAHMAAANH 793  
QY 480 CACCACCAACCCACAGAGACCAACCATCAACCAACCAACCAACCAATCAACAACCTAC 539  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
792 MMWAKYCNAMWACMEHYAAYAIAAAAAAATAAAWMAAHNAHMAAANHMAAANHMM 733  
QY 540 ACTGCTCACCAACA-ACACCAACAGGAATCCAAAACCTCAACAGTCAAATGGAAACCTTCC 598  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
732 MMCAEWAANWMACASCA-CBBHASCMSCANKYHKHHHHWHHCYHHYBHYHBARY 673  
QY 599 ACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACAATCCGAGCACC 658  
Db |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
672 YCYMHBYCWKYCTAAAAAATAAAAKYYYYYAANAYAKYYAWYWKYTMAANRYY 613  
QY 659 CATCAACACCTCATCTC 676  
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

Db 612 AAAAGHYNTYKHYTY 595

RESULT 4  
AG127412

LOCUS AG127412 946 bp DNA linear GSS 04-NOV-2001

DEFINITION Pan troglodytes DNA, clone: PTB-138E06.R, genomic survey sequence.

ACCESSION AG127412

VERSION AG127412.1 GI:16656577

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE BAC end sequences of Library PTB

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 946)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpses@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI.

Location/Qualifiers

1. .946

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

/clone="PTB-138E06.R"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC Library"

505 a 351 c 31 g 23 t 36 others

BASE COUNT

ORIGIN

Query Match 14.6%; Score 104.6; DB 29; Length 946;

Best Local Similarity 46.8%; Pred.No.3.3e-07;

Matches 326; Conservative 0; Mismatches 374; Indels 0; Gaps 0;

Qy 15 AACACATGCATCATCAAGATGCAACAAGCCAGATCAAGAACACACACCCCAACATACCT 74

Db 153 AACAAACACACACACCACCAANACACACACAGGCAACAANNAAAAACCCACACCAAC 212

Qy 75 CACTCAGGATCCTCAGCTTGGATCAGCTTCTCCAACTGTCTGAAATTACATCAAAAC 134

Db 213 AGCAACCAAAAAACACCCAAAAACAAAAACAACCAAGAAACACACACACACAGAAAC 272

Qy 135 CACCACCATACTAGCTTCAACACACACGAGTCAAGTCAAACTGCAACCCACACACACT 194

Db 273 CACCCCAACAAAAAACCAACCAACAAACACCAACCAACCAAGACACCAAAAAA 332

Qy 195 CAAGACTAAACACACAAACCAACCCAAACACAAACCCAGCAGCCCACTACAAAAACAG 254

Db 333 AAAAACTATAACAAACACACAAACCAACCAACCCCAACACCAACCAAAAAAACAACA 392

Qy 255 CCAAAACAAACCAACCAACCAACCCCAATATGATTTCTGCTGAAGTGTTTAACTTTGT 314

Db 393 AAACACACACCAACCAACCAACCCCAACCAACCCCAACCAACCAACCAACCAACCA 452

QY 315 ACCCTGCAGCATATGACGACAAATCCCACTGCTGGGCTATCTGCAAAAGATACCAA 374  
 Db 453 AAAACAACCATGCTGCTCCCAAAACCAACAAACAAAAAACCACCCCAAA 512  
 QY 375 CAAAAACCGAGAAAGAAAACCAACCAACCAAGCTTACAAAAAACCAACCTTCAAGACAA 434  
 Db 513 ACAAAAACAAAAAACCCACACACCAACCAACCAACCAACCAACCAACCAACCAACCA 572  
 QY 435 CAAAAAGATCTTCAAACTTCAAACTTCAAACTTCAAACTTCAAACTTCAAACTTCAAA 494  
 Db 573 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 632  
 QY 495 AGAAGAGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 554  
 Db 633 AAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 692  
 QY 555 CACCACAGAAATCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 614  
 Db 693 AAACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 752  
 QY 615 AGGCAATCTAAGCCTTCTCAAGTCTCAACATCTCGAGACCAACCAACCAACCAACCA 674  
 Db 753 CAGCACCCCAAAAAAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 812  
 QY 675 TCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714  
 Db 813 CCAAAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 852

RESULT 5  
 CNS0073W/c  
 LOCUS  
 DEFINITION  
 Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 BAC14D09 of RPCI-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

AL066784.1 GI:4945247  
 GSS.  
 Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 922)  
 Genoscope.  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see <http://www.fruitfly.org> The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazuo Oosawa and  
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of  
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
 NY. The library is named RPCI-98 and was constructed by partial  
 EcoRI digestion of Drosophila DNA provided by the BDGP from the  
 isogenic strain y2; cn bw ep, the same strain used for the BDGP's  
 P1 and EST libraries. A more detailed description of the library  
 and how to order individual BAC clones, the entire library, or  
 filters for hybridization from the BACPAC Resource Center can be  
 found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
 Location/Qualifiers

FEATURES  
 source

1. .922  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BAC14D09"  
 /clone\_lib="RPCI-98"  
 /note="end : TET3"

BASE COUNT 223 a 95 c 109 g 221 t 274 others

## ORIGIN

Query Match 14.4%; Score 103.2; DB 29; Length 922;  
 Best Local Similarity 24.8%; Pred. No. 5.4e-07;  
 Matches 102; Conservative 164; Mismatches 144; Indels 1; Gaps 1;

QY 181 CAACCCCAACAGCTCAAGACTTAAACAAACACACACACACACACACACACACACACAC 240  
 Db 921 MAMMCMCCCHCCCHCCCHCCCHCCCHCCCHCCCHCCCHCCCHCCCHCCCHCCCHCC 862  
 QY 241 ACTACAAACAAACCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 Db 861 ACAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 802  
 QY 301 GTGTTTAACTTTGTACCTGTCAGCATATGAGAGCAACCAATCAACCTGCTGGGCTAT 360  
 Db 801 ACACWCAAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 743  
 QY 361 AABAGATATCCAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 Db 742 AWAAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 683  
 QY 421 ACCTTCAAGACCAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 Db 682 MWCAAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 623  
 QY 481 ACCACCAAGCCCAACAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 540  
 Db 622 AMWACWCAAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 563  
 QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 591  
 Db 562 AMACWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAMWAM 512

RESULT 6  
 CNS04ENY/c

LOCUS  
 DEFINITION  
 CNS04ENY 919 bp DNA linear GSS 01-SEP-2000  
 Tetraodon nigroviridis genome survey sequence T7 end of clone  
 104P14 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AL287287  
 AL287287.1 GI:8025774  
 GSS; genome survey sequence.  
 Tetraodon nigroviridis  
 Tetraodon nigroviridis

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,  
 Saurin, W. and Weissenbach, J.  
 Estimate of human gene number provided by genome-wide analysis  
 using Tetraodon nigroviridis DNA sequence  
 Nat. Genet. 25 (2), 235-238 (2000)  
 20296633  
 10835645

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
 Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)  
 20359837  
 10899143

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
 Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)  
 20359837  
 10899143

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
 Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)  
 20359837  
 10899143

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)





```

QY 273 CAACACCAATATGATTTTCACTTCGAAAGTGTGTTAACTTTGTACCTCGAGCATATGCG 332
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
961 MAAAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 902
QY 333 CAACAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACAAACCAAGAAAGAA 392
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
901 CMAMMMAAAAMAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMM 842
QY 393 AACACACCAACGCTTACAAAAAACCACCTTCAAGACAAACAAAGAGATCTCAACC 452
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
841 CMCMMAAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 782
QY 453 TCAAAACCACTAAACCAAGGAAGTACCCACCAAGCCCAAGAGAGAGCCCAACCATCAA 512
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
781 CMCMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMM 722
QY 513 CACCACCAAAACAAACATCAACTACATGCTGCTACCAACCAACACC 558
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
721 TCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCC 676

RESULT 10
BX443342/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches
Indels
Gaps
1;
29
753
TACAAAGATGCAACAGCCAGATCAAGAAACACACACCCCAACATACCTCAGGATCCTC 88
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TMMCGGNGMGVNGMGVNGMGVNGMGVNGMGVNGMGVNGMGVNGMGVNGMGVNGMGVNGMG 694

```

```

QY 89 AGCTTGGAAATCAGCTTCTCCAAATCTCTGAAATATATCATCAAAACCCACCATACTAG 148
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
693 MAMMMMMGNGVNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNN 634
QY 149 CTTTCAACAAACACCCAGGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGACTAAACA 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
633 MAMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 574
QY 209 CAACAAACAAACCCCAACCAACCCAGCAGCCACCTACAAACAAACGCGCAAAACCAAC 268
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
573 NMNMCMGMMMMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 514
QY 269 CAAAACAAACCCCAATATGATTTTCACTTCGAAAGTGTGTTAACTTTGACCTCGAGCATAT 328
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
513 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 454
QY 329 GCAGCAACCAATCCAACTGCTGGGTCTCTGCAAAAGATATCTGCAAAAGATATCTGCA 388
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
453 NCANNNCNCAACGCGCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 394
QY 389 AGAAAAACCAACCAACGCTTACAAAAAACCAACCTTCAAGACCAACCAAAAAAGATCTCA 448
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
393 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 334
QY 449 AACCTCAACCACTAAACCAAGAGAGTACCACCAAGCCCAAGAGAGAGAGAGAGAGAG 508
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
333 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 274
QY 509 TCAACACCAACCAACCAACCAACCACTTACACTGCTGCAACCAACCAACCAACCAAGATC 568
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
273 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 214
QY 569 CAAAACTCACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCC 628
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 MMAMCMMAAAACCCCM---MAAACCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCC 157
QY 629 CTTCTCAAGTCTCAACATCCCGAGCAGCCCAATCAACACCTTC 671
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 CCCCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 114

```

```

RESULT 11
BX385076
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
BX385076 1100 bp mRNA linear EST 08-MAY-2003
BX385076 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL006YE12 3-PRIME, mRNA sequence.
BX385076.1 GI:30436528
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1100)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4073.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL006BC06NP1&cluster=4073.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL006BC06NP1.
Location/Qualifiers
1..1100
/organism="Homo sapiens"

```



```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL006YB12"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      433 a 175 c 126 g 196 t 170 others
ORIGIN

```

```

Query Match      13.5%; Score 96.2; DB 13; Length 1100;
Best Local Similarity 38.9%; Pred. No. 6.5e-06;
Matches 167; Conservative 72; Mismatches 190; Indels 0; Gaps 0;

QY 146 TAGCTTCAACACACAGGAGTCAAGTCAAACTCGACCCACACAGTCAAGACTRAAA 205
Db 672 TWWAWWWAAWAAAAAARAAAAAACAACAAAAAAMAAAAAAMAAAAAAMAAAAA 731

QY 206 ACACAACACACACACACACACACACACACACACACACACACACACACACAA 265
Db 732 AAAAAMAAAAAAMWCAAAAAAARAAAAAACAACACACACACACACACACACAA 791

QY 266 CACCAACACACACACACACACACACACACACACACACACACACACACACAA 325
Db 792 AAAMAMCAAAAMCCAMAAAAAAMWMMWMMWMMWMMWMMWMMWMMWMMWMMW 851

QY 326 TATGACGACACACACACACACACACACACACACACACACACACACACACAA 385
Db 852 AAAAAACACCAAMAAAAAARAAAAAAMWMMWMMWMMWMMWMMWMMWMMWMMW 911

QY 386 GAAAGAAAAACACACACACACACACACACACACACACACACACACACACACAA 445
Db 912 MAMWCAAAAAAAMWAAACACACACACACACACACACACACACACACACACAA 971

QY 446 TCAACCTCAACACACACACACACACACACACACACACACACACACACACAA 505
Db 972 MAAAMWCCAMACCCMAAAACACACACACACACACACACACACACACACACAA 1031

QY 506 CCAATCAACACACACACACACACACACACACACACACACACACACACACAA 565
Db 1032 AAAAAACACCAAMWAAAMWAAAMWAAAMWAAAMWAAAMWAAAMWAAAMWAA 1091

QY 566 ATCCAAAAC 574
Db 1092 AAAMWCCMCC 1100

```

```

RESULT 12
CNS01224/c
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence SP6 end of BAC
  BACN08M09 of DrosBAC library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL102202.1
VERSION
  GI:5613813
KEYWORDS
  Drosophila melanogaster (fruit fly)
SOURCE
  Drosophila melanogaster
  Organism
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 1097)
AUTHORS
  Direct Submission
TITLE
  Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
JOURNAL
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (EDGP) -

```

```

http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

```

```

FEATURES             Location/Qualifiers
     source            1..1097
     /organism="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /db_xref="taxon:7227"
     /clone="BACN08M09"
     /clone_lib="DrosBAC"
     /plasmid="pBelobAC11"
     /note="end : SP6"
BASE COUNT      169 a 119 c 196 g 398 t 215 others
ORIGIN

```

```

Query Match      13.4%; Score 95.8; DB 29; Length 1097;
Best Local Similarity 32.7%; Pred. No. 7.6e-06;
Matches 199; Conservative 119; Mismatches 290; Indels 0; Gaps 0;

QY 13 CTAACAATCGAATCATACAAAGATGCAACAGCCAGATCAAGAACACCAACCAATAC 72
Db 1095 MTCMMWMTWMCATNTCMYCMCCMCATACMACMCCCAACMMWCCACCCCAATMCT 1036

QY 73 CTCCTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAAATTTACATCAA 132
Db 1035 CMCAMCMWAAACAAMACACMCMWCAACMACACMACACCCACACCAACACACACACACA 976

QY 133 ACACACACATCTAGCTTCAACAAACACAGGAGTCAAGTCAAACTGGCAACCCACACA 192
Db 975 CMWCCACACAAACACACACACACACACACACACACACACACACACACACACACACAC 916

QY 193 GTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 252
Db 915 AMWACMACACACACACACACACACACACACACACACACACACACACACACACACACAA 856

QY 253 CGCCAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 312
Db 855 AHAACAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 796

QY 313 GTACCTGCGACATATCGAGCAACATCCAACTGCTGGGGTATCTGCAAAAGAAATACCA 372
Db 795 TTTTTTTTTTTTATACWNCACWACWCTYMCATATAATWMCACACACAYMACCATCMYH 736

QY 373 AACAAAAACCCAGGAAGAAACCCACACCAAGCTCAAAAAACCAACCAACCAACCAACCA 432
Db 735 WMTWACMACACACACACACACACACACACACACACACACACACACACACACACACACAA 676

QY 433 ACCAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAAGTACCCACCAACCAAGCCC 492
Db 675 AAAAMCACMMWMAACCCCMCAACACCCCCCCCCCCCCCCCCCCCCACACACACACACAA 616

QY 493 ACAGAGGCAACATCAACACACACACACACACACACACACACACACACACACACACACAA 552
Db 615 MAAAAAAMAAAAAMWCCCAACCMCCCTTMAWACMMYACMMWACMMWACMMWACMMWACMM 556

QY 553 AACACACAGGAATCCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 612
Db 555 TAMCMWACACACACACACACACACACACACACACACACACACACACACACACACACATAC 496

QY 613 GAAGGCAA 620
Db 495 ACAACCMW 488

```

```

RESULT 13
CNS009KS
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence T7 end of BAC #
  BACR19F04 of RPCL-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.

```



```

ACCESSION      AL053801.1      GI:4935176
VERSION        GSS.
KEYWORDS
SOURCE         Drosophila melanogaster (fruit fly)
ORGANISM       Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
               Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 791)
AUTHORS
TITLE          Direct Submission
JOURNAL        Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
               BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
               - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
               collaboration with the Berkeley Drosophila Genome Project (BDGP).
               The BDGP is constructing a physical map of the Drosophila
               melanogaster genome using these BACs. For further information
               please see http://www.fruitfly.org The BDGP Drosophila
               melanogaster BAC library was prepared by Kazutoyo Osoegawa and
               Aaron Mammosser in Pieter de Jong's laboratory in the Department of
               Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
               NY. The library is named RPCI-98 and was constructed by partial
               EcoRI digestion of Drosophila DNA provided by the BDGP from the
               isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
               P1 and EST libraries. A more detailed description of the library
               and how to order individual BAC clones, the entire library, or
               filters for hybridization from the BACPAC Resource Center can be
               found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
               1..791
               /organism="Drosophila melanogaster"
               /mol_type="genomic DNA"
               /db_xref="taxon:7227"
               /clone="BACR19F04"
               /clone_lib="RPCI-98"
               /note="end : T7"
BASE COUNT    464 a 105 c 67 g 78 t 77 others
ORIGIN
Query Match   13.3%; Score 95; DB 29; Length 791;
Best Local Similarity 47.1%; Pred. No. 1.e-05;
Matches 205; Conservative 43; Mismatches 178; Indels 9; Gaps 2;

QY 153 AACACACACGAGTCAAGTCAACCTGCAACCCACACAGCTCAAGCTAAGCAACAC 212
Db 153 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 92
QY 213 AACACCCCAACCAACCCAGCAGCCCACTACAAACAAACGCAACAAACCAACAA 272
Db 93 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 152
QY 273 CAACACCA-ATAATGATTTTCACTTCGAGTGTGTTAACTTTGACCTCGAGCATGCA 331
Db 153 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 212
QY 332 GCACCAATCAACCTGCTGGCTATCTGCAAAAGATACCAACAAACAAACAGGAAGA 391
Db 213 ACWATAAGGACAA-----AAAMAMMAAAAMMAAAAMMAAAAMMAAAAMMA 264
QY 392 AAACACCAACCAAGCTTACAAACAAACCAACCTTCAAGACAAACCAACAAAGATCTCAAC 451
Db 265 AAACMAAAACWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 324
QY 452 CTCAACCACTAAACCAAGGAGTACCCACCAAGCCCAAGCCCAAGAGCCCAACATCA 511
Db 325 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 384
QY 512 ACACCAACCAACCAACATCACTACACTGCTCCACCAACCAACCAACAGGAATCCAA 571
Db 385 AAAAAAAMCWAAMAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 444
QY 572 AAUACAAGTCAAA 586

```

```

Db 445 AACACACAAATATAAA 459
||||| ||||| |||||
CNS00LOO      1101 bp      DNA      linear      GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC;
BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL068607
AL068607.1    GI:4958689
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
               1..1101
               /organism="Drosophila melanogaster"
               /mol_type="genomic DNA"
               /db_xref="taxon:7227"
               /clone="BACR32D23"
               /clone_lib="RPCI-98"
               /note="end : TET3"
BASE COUNT    155 a 166 c 7 g 284 t 489 others
ORIGIN
Query Match   13.3%; Score 95; DB 29; Length 1101;
Best Local Similarity 18.5%; Pred. No. 1.e-05;
Matches 110; Conservative 245; Mismatches 240; Indels 0; Gaps 0;

QY 104 TCTCCAATCTGCTGAATTAATACATCAACAAACACCAACCACTAGCTTCAACACACG 163
Db 455 TCTCCCTCTATTTCTBCCTTTTCCTTCACACACACACACACACACACACACAC 514
QY 164 GAGTCAAGTCAACCTGCAACCCACCAAGTCAAGCTAAGCAACAAACCAACCAAC 223
Db 515 CMHMMWACWMACTMMWCMWTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 574
QY 224 CAACACCCAGCAGCCCACTACAAACAAACGCGCAACAAACCAACCAACCAACCA 283
Db 575 MAYCMYCCMYTTCMCMAHHTWMAACATMAHACHTHTHCMTWTWTWTWTWTWTWT 634
QY 284 ATGATTTTCACTTCGAGTGTTAATCTTTGACCTCGAGCATATATGACGAAACAAT 343
Db 635 MAHTMTWMAACHAHNWTWTWTMMHMAHAYAHCACTWTWTTCAMMMWMAACMMHMM 694
QY 344 CTGCTGGGCTATCTGCAAAAGAAATACAAACAAACAAACCAAGGAAGAAACCAACCA 403
Db 695 MHTMTATCAAMWMAAMMTMTTWTWMAACMRMAMMMCCSCCAMMCMAMMAHMAACMA 754

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2003, 00:01:13 ; Search time 240.957 Seconds  
(without alignments)  
8070.528 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715

Sequence: 1 cacaaagtcacactaacac.....gtagttattataaaaaaaa 715

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1811591 seqs, 1359896290 residues

Total number of hits satisfying chosen parameters: 3623182

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	654.8	91.6	15223	10	US-09-847-173-1
2	445.8	62.3	696	10	US-09-844-645-4
3	300.4	42.0	15225	12	US-09-827-688-10
C 4	85.8	12.0	367378	12	US-10-312-841-2
C 5	83.6	11.7	793	10	US-09-878-574-4304
C 6	82.6	11.6	7758	12	US-10-311-455-1076
C 7	81.4	11.4	16167	12	US-10-311-455-1056
C 8	81.4	11.4	16167	12	US-10-240-485-82
C 9	80	11.2	6668	12	US-10-311-455-1670
C 10	78.8	11.0	1972	12	US-10-017-161-1981
C 11	76.6	10.7	576	10	US-09-878-574-4296
C 12	76.4	10.7	529	10	US-09-983-965-2109
C 13	75.6	10.6	1030	10	US-09-878-574-4306
C 14	75	10.5	577	10	US-09-878-574-4312
C 15	74.8	10.5	6089	12	US-10-311-455-675
C 16	74.8	10.5	11622	12	US-10-311-455-648

C 17	74	10.3	639	10	US-09-878-574-4316	Sequence 4316, Ap
C 18	73.4	10.3	6668	12	US-10-311-455-1669	Sequence 1669, Ap
C 19	72.8	10.2	1083	12	US-10-029-386-20723	Sequence 20723, A
C 20	72.8	10.2	6048	12	US-10-311-455-2004	Sequence 2004, Ap
C 21	72.2	10.1	13606	12	US-10-311-455-1783	Sequence 1783, Ap
C 22	72.2	10.1	13606	12	US-10-240-453-187	Sequence 187, App
C 23	72.2	10.1	13606	14	US-10-239-676-165	Sequence 165, App
C 24	71.8	10.0	2543	12	US-10-029-386-20536	Sequence 20536, A
C 25	71.8	10.0	3163	12	US-10-017-161-1857	Sequence 1857, Ap
C 26	71.8	10.0	14006	12	US-10-311-455-1931	Sequence 1931, Ap
C 27	71.4	10.0	1075	9	US-09-864-761-19241	Sequence 19241, A
C 28	71.4	10.0	1403	9	US-09-864-761-19241	Sequence 2513, Ap
C 29	71.2	10.0	1635	9	US-09-864-761-20241	Sequence 20241, A
C 30	71.2	10.0	1973	9	US-09-864-761-3471	Sequence 3471, Ap
C 31	71	9.9	13732	12	US-10-311-455-1793	Sequence 1793, Ap
C 32	70.8	9.9	11996	12	US-10-240-485-45	Sequence 45, Appl
C 33	70.6	9.9	8147	12	US-10-311-455-401	Sequence 401, Appl
C 34	70.6	9.9	9415	12	US-10-311-455-268	Sequence 268, Appl
C 35	70.4	9.8	6270	12	US-10-311-455-378	Sequence 378, Appl
C 36	70.4	9.8	6270	12	US-10-240-453-20	Sequence 20, Appl
C 37	70.4	9.8	6270	14	US-10-239-676-12	Sequence 12, Appl
C 38	70.4	9.8	9539	12	US-10-240-453-54	Sequence 54, Appl
C 39	70.4	9.8	9539	14	US-10-239-676-52	Sequence 52, Appl
C 40	70.2	9.8	10467	12	US-10-240-453-327	Sequence 327, Appl
C 41	70.2	9.8	17869	12	US-10-311-455-78	Sequence 78, Appl
C 42	70	9.8	576	9	US-09-864-761-26582	Sequence 26582, A
C 43	70	9.8	6065	12	US-10-311-455-477	Sequence 477, App
C 44	69.6	9.7	5774	12	US-10-311-455-1136	Sequence 1136, Ap
C 45	69.6	9.7	9219	12	US-10-311-455-2400	Sequence 2400, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-847-173-1

; Sequence 1, Application US/09847173

; Publication No. US2002018228A1

; GENERAL INFORMATION:

; APPLICANT: Collins, Peter L.

; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY

; SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/847,173

; FILING DATE: 03-May-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/720,132

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Parmelee, Steven W.

; REGISTRATION NUMBER: 31,990

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-467-9600

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15223 base pairs

; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

Query Match      91.6%; Score 654.8; DB 10; Length 15223;
Best Local Similarity 94.8%; Pred. No. 4e-164;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 60
DB 4887 CACAAAGTCACACCAACCACTGCAATCATACAAAGATGCAACAAAGCCAGATCAAGAACACA 4946

QY 61 ACCCCAAATACCTCACTCAGATCCTCAGCTTGGAATCAGCTTCTCAATCTGTCTGAA 120
DB 4947 ACCCCAAATACCTCACTCAGATCCTCAGCTTGGAATCAGCTTCTCAATCTGTCTGAA 5006

QY 121 ATTACATCAAAACACCAACCATATAGCTTCAACAAACAGGAGTCAAGTCAACCTG 180
DB 5007 ATTACATCAAAACACCAACCATATAGCTTCAACAAACAGGAGTCAAGTCAACCTG 5066

QY 181 CAACCCCAACAGTCAGACTAAACACCAACAAACCCCAACCAACCAACCAACCAACCA 240
DB 5067 CAATCCCAACAGTCAGACTAAACACCAACAAACCCCAACCAACCAACCAACCAACCA 5126

QY 241 ACTACAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 5127 ACCCAAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186

QY 301 GTGTTTAACTTTGTACCTGAGCATATGCGAGCAATCAACCTGCTGGGCTATCTGC 360
DB 5187 GTGTTTAACTTTGTACCTGAGCATATGCGAGCAATCAACCTGCTGGGCTATCTGC 5246

QY 361 AAAAGAAATACCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 5247 AAAAGAAATACCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAGACACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 5307 ACCCTCAAGACACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCCAAGCCACACAGAGAGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 5367 ACCCAAGCCACACAGAGAGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 5427 CTACTCACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5486

QY 601 TCAACCTCTCGAAGGCAATTAAGCCCTTCTCAAGTCTTCAACCAACCAACCAACCAACCAACCA 660
DB 5487 TCAACCTCTCGAAGGCAATTAAGCCCTTCTCAAGTCTTCAACCAACCAACCAACCAACCAACCA 5546

QY 661 TCACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
DB 5547 TCACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5600
```

## RESULT 2

```
US-09-844-645-4
; Sequence 4, Application US/09844645
; Patent No. US20020102242A1
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
;             McDaniel, Larry S.
;             Curiel, David T.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR
; ADMINISTERING PNEUMOCOCCAL DNA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
```

```
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/844,645
; FILING DATE: 27-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,505
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-844-645-4

Query Match      62.3%; Score 445.8; DB 10; Length 696;
Best Local Similarity 95.3%; Pred. No. 5.8e-109;
Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1 CACAAAGTCACACTAACCACTGCATCATACAGATGCAACAGCCAGATCAAGAACACA 60
DB 199 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 258

QY 61 ACCCCAAATACCTCACTCAGATCCTCAGCTTGGAATCAGCTTCTCAATCTGTCTGAA 120
DB 259 ACCCCAAATACCTCACTCAGATCCTCAGCTTGGAATCAGCTTCTCAATCTGTCTGAA 318

QY 121 ATTACATCAAAACCAACCAACCATATAGCTTCAACCAACCAACCAACCAACCAACCAACCA 180
DB 319 ATTACATCAAAACCAACCAACCATATAGCTTCAACCAACCAACCAACCAACCAACCAACCA 378

QY 181 CAACCCCAACAGTCAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 239
DB 379 CAATCCCAACAGTCAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 438

QY 240 CACTTCAAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 299
DB 439 CACCAAAACCAACGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498

QY 300 AGTGTTTAACTTTGTACCTGCGAGCATATGCGAGCAACCAATCCAACTCTGCTGGGCTATCTG 359
DB 499 AGTGTTCACCTTTGTACCTGCGAGCATATGCGAGCAACCAATCCAACTCTGCTGGGCTATCTG 558

QY 360 CAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 419
DB 559 CAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618

QY 420 AACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 479
DB 619 AACCTTCAAGACCAACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 678

QY 480 CACCACCAAGCCC 492
DB 679 CACCACCAAGCCC 691
```

## RESULT 3

US-09-827-688-10  
; Sequence 10, Application US/09827688  
; Publication No. US20030165476A1  
; GENERAL INFORMATION:  
; APPLICANT: ORSON, FRANK  
; APPLICANT: KINSEY, BERNIA  
; APPLICANT: BHOAGAL, BALBIR  
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION  
; FILE REFERENCE: P01949US/10004014  
; CURRENT APPLICATION NUMBER: US/09/827,688  
; CURRENT FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 2001-04-06  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 15225  
; TYPE: DNA  
; ORGANISM: RSV  
US-09-827-688-10

Query Match 42.0%; Score 300.4; DB 12; Length 15225;  
Best Local Similarity 64.6%; Pred. No. 1.5e-69;  
Matches 464; Conservative 0; Mismatches 251; Indels 3; Gaps 1;

QY	1	CACAAAGTTCACACTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA	60
DB	4888	CACAAAGTTCACACTAACAGGTTCACAGTTCACACATATAAAACACACTGAAAAAAC	4947
QY	61	ACCCCAACATCTCTACTCAGGATCTCTAGCTTTGGAATCAGCTTTCTCCATCTGTCTGAA	120
DB	4948	ATCACCACCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT	5007
QY	121	ATTACATCAACACACACACCTACTACTTCTTCAACACACAGGAGTCAAGTCAACCTTG	180
DB	5008	ACCACATTCACCAATTCACACAAATTCAGCCACACATCTACCCACACAAAGTCAGAAACA	5067
QY	181	CAACCCACACAGCTCAAGACTTAAACACACAAACACACACACACACACACACACACAC	240
DB	5068	CACCAAC	5127
QY	241	ACTACAAAC	300
DB	5128	AGCACAAAC	5187
QY	301	GTGTTTAACTTTGACCTTGCAGCATATGCAGCAACATCCACCTGCTGGGCTATCTGC	360
DB	5188	GTGTTTAACTTTGACCTTGCAGCATATGCAGCAACATCCACCTGCTGGGCTATCTGC	5247
QY	361	AAAAGAATACCAAC	420
DB	5248	AAAAGAATACCAAC	5307
QY	421	ACCTTCAAG---ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	477
DB	5308	ACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	5367
QY	478	CCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	537
DB	5368	ACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	5427
QY	538	ACACTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	597
DB	5428	TCACATCCATGCTGAC	5487
QY	598	CACTCAACTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACACATCCCGAGCAC	657
DB	5488	CACTCAACTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACACATCCCGAGCAC	5547
QY	658	CCATCAACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	715

Db 5548 TCTACATCAAAATCCACCCAAATACCAATCACATGCTTAGTTATTCAAAAACACTACA 5605

## RESULT 4

US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2

Query Match 12.0%; Score 85.8; DB 12; Length 3673778;  
Best Local Similarity 47.0%; Pred. No. 3.1e-11;  
Matches 334; Conservative 0; Mismatches 372; Indels 5; Gaps 2;

QY	9	CACACTAACCAACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACACACACCCCAAC	68
DB	2435804	CTCAACAAACCTACAAACACGAAACAAAAAATCCACTAACATATAAAACACACACACC	2435745
QY	69	ATACCTCACTCAGGATCTCTCAGCTTGGATCAGCTTCTCCATCTGTCTGAAATTCATC	128
DB	2435744	ATCCCGCAACAAACCTTACAAAAATAAAAAAACAACCAATATAAAACACACACCATC	2435685
QY	129	ACAAACCCACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTGCAACCCAC	188
DB	2435684	CCACACAAACCTTACAAAAATAAAAAAATAACCAACGAAAAACCACTATTCC	2435625
QY	189	AACAGTCAAGACTAAAAACACAAACACACACACACACACACACACACACACACACAC	248
DB	2435624	AACAAACCTACAAAAATAAAAAAATAACCAACCAATATAAAACACATATCCCAAC	2435565
QY	249	ACAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	308
DB	2435564	AAAC-CTACAAAAATAAAACAAAAACCCCAATTTACCAATAAAAAACCACTATCCC	2435506
QY	309	CTTTGTACCTGCAGCATATGCAGCAACATCCACCTGCTGGGCTATCTGAAAGAAAT	368
DB	2435505	CAACAAACCTTACAAAAACGAAAAAACCACCACTAACCAATAAAAAACACACACTAT	2435446
QY	369	ACCAACAAAA---AACCGAAAAAACCACCAACCAACCAACCAACCAACCAACCAACCA	424
DB	2435445	CCCAACCAACCTTACAAAAATAAAAAAATAACCAACCAATATAAAACACCACTATCC	2435386
QY	425	TCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	484
DB	2435385	CAACAAACCTTACAAAAATAAAAAAATAACCAACCAATATAAAACACCACTATCCC	2435326
QY	485	CCAGCCCAACGAGGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAACCA	544
DB	2435325	CACACCTTACAAAAATAAAAAAATAACCAACCAATATAAAAAACCACTATCCC	2435266
QY	545	TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	604
DB	2435265	TACAAACCTTACAAAAATAAAAAAATAACCAACCAATATAAAACCACTATCCC	2435206
QY	605	CCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCCATCAC	664
DB	2435205	CCTCAACAAACCTTACAAAAACGCAACCAACCAACCAACCAACCAACCAACCAACCA	2435146



DB 2057 ACTCCACACACAAAAAACAACTAAAACTCTCTAAACACAAAAAAC 2005

RESULT 7

US-10-311-455-1056/c

; Sequence 1056, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIENPEROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10311,455

; PRIOR FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1056

; LENGTH: 16167

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732

; OTHER INFORMATION: n is a o r c o r t

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 5837..5838

; OTHER INFORMATION: n is a o r c o r t

US-10-311-455-1056

Query Match	11.4%;	Score 81.4;	DB 12;	Length 16167;
Best Local Similarity	46.6%;	Pred. No. 4e-11;		
Matches 235;	Conservative 0;	Mismatches 269;	Indels 0;	Gaps 0;
QY	109	AATCTGCTGHAATTACATCAACCAACCCACCATCTAGCTTCAACAACACCGAGATC	168	
Db	5866	ATTCTTACTTTTCTTCACACCAACCAACGNNNAACCAACGCAAAAAAAAAAAAAAAAAAAAA	5807	
QY	169	AAGTCAAACCTGCAACCCCAACACAGTCAAGACTTAAABACCAACCAACCCCAACACAA	228	
Db	5806	AA	5747	
QY	229	CCGACGACGCCACTACAAACCAACGCGCAACCAACCAACCAACCAACCCCAATATGAT	288	
Db	5746	AAAAAAAAAAAAAGNAAAAAAAAAAAAACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACA	5687	
QY	289	TTTCACCTCGAAGTGTTTAACTTTGTACCTGACGATATGCAGCAACAATCCAACCTGC	348	
Db	5686	AA	5627	
QY	349	TGGGCTATCTGCAAAAGCAATACCAACCAAAAAAAAAACGAGGAAGAAAAACCAACCAAGCCT	408	
Db	5626	CGAA	5567	
QY	409	ACAAAAAAAAACCACTTCAAGACAACCAAAAAAAAAAGATCTCAAACTCAAAACCACTTAAACCA	468	
Db	5566	AA	5507	
QY	469	AAGGAAGTACCCACCAACGACCCCAAGAGAGCGCAACCACTCAACCAACCAACCAACAA	528	
Db	5506	AACCTTTTACCAAAACCCCACTTACATCACTTACCTTAAACCTTCTACCAAAACCTAA	5447	
QY	529	ATCACAACTACTGTGTCAACAACAACCAACGAGGAATCAAAACACTCAAGTCAAAATG	588	

```

Db      5446 TCTCCCCAAATTCCTCACAACCTTCTTATCACACACCTTAATAAATAAAAAAAAAAAAAA 5387
QY      589 GAACCTTCCACTCAACCTCTCTCC 612
Db      5386 AAAAAAAAAAATAAAATCTCTCTCC 5363

RESULT 8
US-10-240-485-82/c
; Sequence 82, Application US/10240485
; Publication No: US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 82
; LENGTH: 16167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4003, 5529, 5534, 5553, 5570, 5573, 5577..5578, 5592, 5732)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5837..)
; US-10-240-485-82

```

Query Match	11.4%;	Score 81.4;	DB 12;	Length 16167;
Best Local Similarity	46.6%;	pred. No. 4e-11;		
Matches 235;	Conservative 0;	Mismatches 269;	Indels 0;	Gaps 0;
Qy	109	AATCTGTCTGAAATTATCATCAAAACCCACCAACATACTAGCTTCTCAACAACACACAGGAGTC	168	
Db	5866	ATTCTTACTTTTCTCACACCACAAACGNNNAACAAACGAAAAAATAAAAAAATAAAAAA	5807	
Qy	169	AGTCTAAACCTGCACCCCAACAGCTCAAGACTTAAAAACAACAACACCCCAACACAA	228	
Db	5806	AAACGAAAAACAA	5747	
Qy	229	CCGAGCAGCCCTACAAAAACACGCCAAAAACAACCCACCAACCAACCTAATATGAT	288	
Db	5746	AAAAAAAAAAAAACGAAAAAATAAAAAACGAAAAAATAAAAAACCAAAAAAATAACA	5687	
Qy	289	TTTCACCTCGAAGTGTTTAACTTTGTACCTCGACGATATGCGACAACTCAACCTGC	348	
Db	5686	AACG	5627	
Qy	349	TGGGCTATCTGCAAAAGATACCAACAATAAACACGAGAAAGAAAACCAACCAAGCCT	408	
Db	5626	CGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCNNCNAANAATAA	5567	
Qy	409	ACAAAAAACCACTTCAGACCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCA	468	
Db	5566	AAAAAAAAAAAAAATAACCAAAAAAATAAAAAAATAAAAAAATACGAAACCCCAACCCCT	5507	







## RESULT 13

US-09-878-574-4306/c  
; Sequence 4306, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Thompson, Michael D.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1998-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4306  
; LENGTH: 1030  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1030)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-D11  
US-09-878-574-4306

Query Match 10.6%; Score 75.6; DB 10; Length 1030;  
Best Local Similarity 41.0%; Pred. No. 4.1e-10;  
Matches 283; Conservative 0; Mismatches 397; Indels 11; Gaps 2;  
QY 1 CACAAAGTCACACTAACCTCACTCAGGATCCTCAGTTGGAATCAGCTTCTCCAAATCTGCTGAA 60  
DB 789 CACAAC 730  
QY 61 ACCCAACACATCTCCTCAGGATCCTCAGTTGGAATCAGCTTCTCCAAATCTGCTGAA 120  
DB 729 AACAAAC 671  
QY 121 ATTACATCAAAAC 180  
DB 670 -----NACAAAC 618  
QY 181 CAACCCAC 240  
DB 617 CANANAC 558  
QY 241 ACTACAAAC 300  
DB 557 ANCNANANAC 498  
QY 301 GTGTTTAACTTTGACCTTGAGCATATGACGACATATGACGACATATGACGCTATCTGC 360  
DB 497 CCCCCAC 438  
QY 361 AAAAGAATACCAAC 420  
DB 437 ACNAC 378  
QY 421 ACCTTCAAGAC 480  
DB 377 ACAAAC 321  
QY 481 ACCCAAGCCAC 540  
DB 320 ANAAACANAC 261  
QY 541 CTGCTTCAAC 600  
DB 260 CNACNAAAGCNAC 201  
QY 601 TCAACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCCACAAACATCCGAGCACCA 660

## RESULT 14

US-09-878-574-4312/c  
; Sequence 4312, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Thompson, Michael D.  
; APPLICANT: La Rosa, Thomas J.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 4312  
; LENGTH: 577  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(577)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-G11  
US-09-878-574-4312

Query Match 10.5%; Score 75; DB 10; Length 577;  
Best Local Similarity 50.0%; Pred. No. 4.6e-10;  
Matches 232; Conservative 0; Mismatches 230; Indels 2; Gaps 2;  
QY 131 AAACACACACATCTAGTCTTCAACACACACACACACACACACACACACACACACAC 190  
DB 533 AAAC 474  
QY 191 CAGTCAAGACTTAAAC 250  
DB 473 AAC 414  
QY 251 AACCCCAAC 309  
DB 413 CACAAC 354  
QY 310 TTTGTACCTGACGATATGAGCAACATCTCAACCTGTGGGTATCTGCAAAAGATA 369  
DB 353 CAAAAACCCCAAC 294  
QY 370 CCAAC 428  
DB 293 CCAAC 234  
QY 429 GAC 488  
DB 233 AANAAACCCCAAC 174  
QY 489 GCCCAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 548  
DB 173 CACCCCAAC 114  
QY 549 CAAAC 592  
DB 113 CAAAC 70

## RESULT 15

US-10-311-455-675/c

Search completed: October 30, 2003, 03:23:54  
Job time : 246.957 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 21:37:29 ; Search time 55.5382 Seconds  
(without alignments)  
5682.373 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaaagtcacactaacac.....gtagtattataaaaaaaaaa 715

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	920	2	US-08-467-963C-7
2	715	100.0	920	2	US-08-838-189D-7
3	715	100.0	920	3	US-08-852-344D-7
4	715	100.0	920	3	US-08-344-639E-7
5	715	100.0	920	3	US-08-467-969A-7
6	715	100.0	920	3	US-08-467-961A-7
7	715	100.0	920	3	US-08-001-554A-7
8	696	97.3	894	2	US-08-467-963C-28
9	696	97.3	894	2	US-08-838-189D-28
10	696	97.3	894	3	US-08-852-344D-28
11	696	97.3	894	3	US-08-344-639E-28
12	654.8	91.6	15222	2	US-08-801-898A-23
13	654.8	91.6	15222	2	US-08-962-690-12
14	654.8	91.6	15223	2	US-08-832-403A-1
15	654.8	91.6	15223	3	US-08-720-132-1
16	300.4	42.0	15225	2	US-08-892-403A-2
17	212.8	29.8	1050	4	US-08-836-501-76
18	212.8	29.8	1050	4	US-09-626-830-76
19	210.8	29.5	303	3	US-08-721-979A-1
20	210.8	29.5	303	3	US-08-836-504A-1
21	210.8	29.5	303	3	US-08-836-501-1
22	210.8	29.5	303	4	US-09-654-289-1
23	210.8	29.5	303	4	US-09-582-876-1
24	210.8	29.5	303	4	US-09-626-830-1
25	209.6	29.3	1071	3	US-08-836-501-77
26	209.6	29.3	1071	4	US-08-626-830-77
27	207.6	29.0	303	3	US-08-721-979A-3

28 207.6 29.0 303 3 US-08-836-504A-3  
29 207.6 29.0 303 3 US-08-836-501-3  
30 207.6 29.0 303 4 US-09-654-289-3  
31 207.6 29.0 303 4 US-09-582-876-3  
32 207.6 29.0 303 4 US-09-626-830-3  
33 207.6 29.0 303 4 US-09-626-830-3  
34 201.2 28.1 303 3 US-08-721-979A-14  
35 201.2 28.1 303 3 US-08-836-501-14  
36 201.2 28.1 303 4 US-09-654-289-14  
37 201.2 28.1 303 4 US-09-582-876-14  
38 128.6 18.0 183 3 US-09-626-830-14  
39 128.6 18.0 183 3 US-08-721-979A-23  
40 128.6 18.0 183 3 US-08-836-501-23  
41 128.6 18.0 183 4 US-09-654-289-23  
42 128.6 18.0 183 4 US-09-582-876-23  
43 126 17.6 303 3 US-09-626-830-23  
44 126 17.6 303 3 US-08-721-979A-2  
45 126 17.6 303 3 US-08-836-504A-2  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-08-467-963C-7

; Sequence 7, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 920 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-467-963C-7

```
Query Match      100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTCAATCATACAGATGCAACAGGCGAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTAACTCAATCATACAGATGCAACAGGCGAGATCAAGAACACA 265

QY 61 ACCCAAACATCTCACTCAGGATCTCAGCTTGAATCAGCTTCTCCAACTGTCTGAA 120
DB 266 ACCCAAACATCTCACTCAGGATCTCAGCTTGAATCAGCTTCTCCAACTGTCTGAA 325

QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 385

QY 181 CAACCCCAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCA 240
DB 386 CAACCCCAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCA 445

QY 241 ACTCAAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTCAAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 360

QY 301 GTGTTTAACTTTGTACCTCGCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 506 GTGTTTAACTTTGTACCTCGCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 565

QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685

QY 481 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 686 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 745

QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805

QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGCACCCA 660
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAATCCGAGCACCCA 865

QY 661 TCACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
```

```
RESULT 2
US-08-838-189D-7
; Sequence 7, Application US/08038189D
; Patent No. 598169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: ENASHYV, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-838-189D-7
```

```
Query Match      100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTCAATCATACAGATGCAACAGGCGAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTAACTCAATCATACAGATGCAACAGGCGAGATCAAGAACACA 265

QY 61 ACCCAAACATCTCACTCAGGATCTCAGCTTGAATCAGCTTCTCCAACTGTCTGAA 120
DB 266 ACCCAAACATCTCACTCAGGATCTCAGCTTGAATCAGCTTCTCCAACTGTCTGAA 325

QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 385

QY 181 CAACCCCAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCA 240
DB 386 CAACCCCAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAACCAACCA 445

QY 241 ACTCAAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTCAAAACCAACCGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 360

QY 301 GTGTTTAACTTTGTACCTCGCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 506 GTGTTTAACTTTGTACCTCGCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 565

QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685

QY 481 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 686 ACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 745

QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 746 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805
```



```
/ FILING DATE: 06-JAN-1992
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-391 MIS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ TELEX: 065-24567 SIMBAS
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 920 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-344-639E-7

Query Match
Best Local Similarity 100.0%; Score 715; DB 3; Length 920;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCCAGATCAAGAACACA 60
DB CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCCAGATCAAGAACACA 265
QY 61 ACCCCAACTACCTCAGTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 120
DB ACCCCAACTACCTCAGTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 325
QY 121 ATTACATCAAAACCCACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 180
DB ATTACATCAAAACCCACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 385
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACACACACACACACACACACACACAC 240
DB CAACCCCAACAGTCAAGACTTAAACCAACACACACACACACACACACACACAC 445
QY 241 ACTACAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB ACTACAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505
QY 301 GTGTTTAACTTTGTACCTGAGCATATGACGACCAATCCAACTGCTGGCTATCTGC 360
DB GTGTTTAACTTTGTACCTGAGCATATGACGACCAATCCAACTGCTGGCTATCTGC 565
QY 361 AAAGAAATACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420
DB AAAGAAATACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 625
QY 421 ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
DB ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 685
QY 481 ACCCAAGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 540
DB ACCCAAGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 745
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
DB CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805
QY 601 TCAACCTCTCCGAGGCAATCTAGCCCTTCTCAGCTTCCAGCAACATCCGAGCACCA 660
DB TCAACCTCTCCGAGGCAATCTAGCCCTTCTCAGCTTCCAGCAACATCCGAGCACCA 865
QY 661 TCACAAACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB TCACAAACCTCTATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
```

RESULT 5

```
US-08-467-969A-7
; Sequence 7, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasysbyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1K7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-467-969A-7
```

```
Query Match 100.0%; Score 715; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCCAGATCAAGAACACA 60
DB CACAAAGTCACACTAACCTGCAATCATACAAAGTCAAGCCAGATCAAGAACACA 265
QY 61 ACCCCAACTACCTCAGTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 120
DB ACCCCAACTACCTCAGTCAGGATCCTCAGCTTGAATCAGCTTCCAAATCTGTGAA 325
QY 121 ATTACATCAAAACCCACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 180
DB ATTACATCAAAACCCACCATCTAGCTTCAACACACACAGGAGTCAAGTCAAACTG 385
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACACACACACACACACACACACACAC 240
DB CAACCCCAACAGTCAAGACTTAAACCAACACACACACACACACACACACACAC 445
QY 241 ACTACAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB ACTACAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505
```





```
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001,554A
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-286
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-001-554A-7

Query Match 100.0%; Score 715; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAAAGATCAACAAAGCCAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTAACTGCAATCATACAAAGATCAACAAAGCCAGATCAAGAACACA 265
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAATCAGTCTTCAATCTCTGTGAA 120
DB 266 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAATCAGTCTTCAATCTCTGTGAA 325
QY 121 ATTACATCAAAACACCAACCATCTAGCTTCAACACACAGCAGGATCAAGTCAAACTG 180
DB 326 ATTACATCAAAACACCAACCATCTAGCTTCAACACACAGGATCAAGTCAAACTG 385
QY 181 CAACCCCAACAGTCAAGCTAAAGATCAACAAACAAACCCCAACCAACCCAGCAAGCCC 240
DB 386 CAACCCCAACAGTCAAGCTAAAGATCAACAAACAAACCCCAACCAACCCAGCAAGCCC 445
QY 241 ACTCAAAAGACGCGAAACAAACACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 300
DB 446 ACTCAAAAGACGCGAAACAAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 505
QY 301 GTGTTTAACTTTGTACCTGAGCATATGACAGCAACATCCAACTCTGCTGGCTATCTGC 360
DB 506 GTGTTTAACTTTGTACCTGAGCATATGACAGCAACATCCAACTCTGCTGGCTATCTGC 565
QY 361 AAAAGAAATACAAACAAACAAACCCAGGAAGAAACCAACCAACCAAGCTTACAAAAACCA 420
DB 566 AAAAGAAATACAAACAAACAAACCCAGGAAGAAACCAACCAACCAAGCTTACAAAAACCA 625
QY 421 ACCTTCAAGACACCAAAAGATCTCAACCTCAACCACTAAACCAACCAAGGAAGTACCC 480
DB 626 ACCTTCAAGACACCAAAAGATCTCAACCTCAACCACTAAACCAACCAAGGAAGTACCC 685
QY 481 ACCCAAGGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCACTCACTACA 540
DB 686 ACCCAAGGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCACTCACTACA 745
QY 541 CTGCTCACCACCAACCAAGGAATCCAAACCTCAACAGTCAAAATGGAACCTTCTCAC 600
DB 541 CTGCTCACCACCAACCAAGGAATCCAAACCTCAACAGTCAAAATGGAACCTTCTCAC 600

; 746 CTGCTCACCACCAACCAAGGAATCCAAACCTCAACAGTCAAAATGGAACCTTCTCAC 805
; 601 TCAACCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGCACCCA 660
; 806 TCAACCTCTCCGAAGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGCACCCA 865
; 661 TCACACCCCTCATCTCCACCCCAACCAACAGCGCAGTAGTATTAAAAA 715
; 866 TCACACCCCTCATCTCCACCCCAACCAACAGCGCAGTAGTATTAAAAA 920

RESULT 8
US-08-467-963C-28
; Sequence 28, Application US/08467963C
; Patent No. 5968776
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-963C-28

Query Match 97.3%; Score 696; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.8e-180;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAAAGATCAACAAAGCCAGATCAAGAACACA 60
DB 199 CACAAAGTCACACTAACTGCAATCATACAAAGATCAACAAAGCCAGATCAAGAACACA 258
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAATCAGTCTTCCATCTGTCTGAA 120
DB 259 ACCCAACATACCTCACTCAGGATCCTCAGTTGGAATCAGTCTTCCATCTGTCTGAA 318
```

121 ATTACATCAAAACCCACCACTACTAGTCTTCAACACACAGGAGTCAAGTCAAACTG 180  
122 |||||  
319 ATTACATCAAAACCCACCACTACTAGTCTTCAACACACAGGAGTCAAGTCAAACTG 378  
320 |||||  
181 CAACCCCAACAGTCAAGACTTAAACCAACAAACAAACCAACCAACCAACCAACCA 240  
182 |||||  
379 CAACCCCAACAGTCAAGACTTAAACCAACAAACCAACCAACCAACCAACCAACCA 438  
380 |||||  
241 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
242 |||||  
439 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
440 |||||  
301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 360  
302 |||||  
499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 558  
500 |||||  
361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
362 |||||  
559 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
560 |||||  
421 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
422 |||||  
619 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
620 |||||  
481 ACCACCAAGCCCAACAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
482 |||||  
679 ACCACCAAGCCCAACAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
680 |||||  
541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
542 |||||  
739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
740 |||||  
601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 660  
602 |||||  
799 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 858  
800 |||||  
661 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 696  
662 |||||  
859 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCA 894  
860 |||||

## RESULT 9

US-08-838-189D-28  
Sequence 28, Application US/08838189D  
Patent No. 5998169  
GENERAL INFORMATION:  
APPLICANT: KLEIN, Michel H  
APPLICANT: DU, Run-Pan  
APPLICANT: EWASZYHN, Mary E  
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A  
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST  
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
City: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1K7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,189D  
FILING DATE: 16-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/001,554  
FILING DATE: 06-JAN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9200117.1  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jlb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 894 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-838-189D-28

Query Match 97.3%; Score 696; DB 2; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.8e-180;  
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACA 60  
DB 199 CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACA 258  
QY 61 ACCCCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTCTGAA 120  
DB 259 ACCCCCAACATACCTCACTCAGGATCCTCAGCTTGGAAATCAGCTTCTCCAAATCTCTGAA 318  
QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180  
DB 319 ATTACATCAACAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAACTG 378  
QY 181 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240  
DB 379 CAACCCCAACAGTCAAGACTTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 438  
QY 241 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 439 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCAACCA 360  
DB 499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCAACCA 558  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 559 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
QY 421 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 619 ACCTTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
QY 481 ACCACCAAGCCCAACAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
DB 679 ACCACCAAGCCCAACAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
DB 739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
QY 601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCAACCA 660  
DB 799 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCAACCA 858  
QY 661 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCAACCA 696  
DB 859 TCACACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGCTTCCCAACCAACCAACCAACCA 894

## RESULT 10

US-08-852-344D-28



```
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-344-639E-28

Query Match      97.3%; Score 696; DB 3; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.8e-180;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAAGCAAGTCAGATCATCAAGATGCAAGCAAGCCAGATCAAGACACA 60
Db 199 CACAAGTCACACTAAGCAAGTCAGATCATCAAGATGCAAGCAAGCCAGATCAAGACACA 258
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGATCAGTTCTCCAACTCTCTGAA 120
Db 259 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGATCAGTTCTCCAACTCTCTGAA 318
QY 121 ATTACATCAAAACCAACCATACATAGTCTGATCTCAACACACAGGAGTCAAGTCAAACTG 180
Db 319 ATTACATCAAAACCAACCATACATAGTCTGATCTCAACACACAGGAGTCAAGTCAAACTG 378
QY 181 CAACCCCAACAGTCAAGATCTTAAACACACACACACACACACACACACACACACACAC 240
Db 379 CAACCCCAACAGTCAAGATCTTAAACACACACACACACACACACACACACACACACAC 438
QY 241 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 439 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498
QY 301 GTGTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTCTGGCTATCTGC 360
Db 499 GTGTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTCTGGCTATCTGC 558
QY 361 AAAAGATATACCAACCAACCAAGGAAAGCAAGCAACCAACCAACCAACCAACCAACCA 420
Db 559 AAAAGATATACCAACCAACCAAGGAAAGCAAGCAACCAACCAACCAACCAACCAACCA 618
QY 421 ACCTTCAAGCAACCAACCAAGATCTCAACCTCAACCAACCAACCAACCAACCAAGGATACCC 480
Db 619 ACCTTCAAGCAACCAACCAAGATCTCAACCTCAACCAACCAACCAACCAACCAAGGATACCC 678
QY 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
Db 679 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738
QY 541 CTGCTCACCACCAACCAACCAAGGAAATCCAAACTCAAGTCAAGTCAAGTCAAGTCAAGT 600
Db 739 CTGCTCACCACCAACCAACCAAGGAAATCCAAACTCAAGTCAAGTCAAGTCAAGTCAAGT 798
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGACCCCA 660
Db 799 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACATCCGAGACCCCA 858
QY 661 TCACACCTCTATCTCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 696
Db 859 TCACACCTCTATCTCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 894

RESULT 12
US-08-801-898A-23
; Sequence 23, Application US/08801898A
; Patent No. 5998602
; GENERAL INFORMATION:
; APPLICANT: Torrence, Paul F.
; APPLICANT: Silverman, Robert H.
; APPLICANT: Cirino, Nick M.
; APPLICANT: Li, Guiying
; APPLICANT: Xiao, Wei
; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE
; NUMBER OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,898A
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8656-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15222 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: RSV-A2
; LOCATION: 1...15222
; OTHER INFORMATION:
; US-08-801-898A-23

Query Match      91.6%; Score 654.8; DB 2; Length 15222;
Best Local Similarity 94.8%; Pred. No. 1.5e-168;
Matches 677; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAAGCAAGTCAGATCATCAAGATGCAAGCAAGCCAGATCAAGACACA 60
Db 4886 CACAAGTCACACTAAGCAAGTCAGATCATCAAGATGCAAGCAAGCCAGATCAAGACACA 4945
QY 61 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGATCAGTTCTCCAACTCTCTGAA 120
Db 4946 ACCCAACATACCTCACTCAGGATCCTCAGCTTGGATCAGTTCTCCAACTCTCTGAA 5005
QY 121 ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 180
Db 5006 ATTACATCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5065
QY 181 CAACCCCAACAGTCAAGATCTTAAACACACACACCAACCAACCAACCAACCAACCAACCA 240
Db 5066 CAATCCCAACAGTCAAGATCTTAAACACACACACCAACCAACCAACCAACCAACCAACCA 5125
QY 241 ACTACAAACCAAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5126 ACCCAACCAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185
QY 301 GTGTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTCTGGCTATCTGC 360
Db 5186 GTGTTAACTTTGTACCTGAGCATATGCAAGCAACCAATCCAACTCTGGCTATCTGC 5245
QY 361 AAAAGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 5246 AAAAGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5305
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
Db 5306 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5365
QY 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
```



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-892-403A-1

Query Match          91.6%; Score 654.8; DB 2; Length 15223;
Best Local Similarity 94.8%; Pred. No. 1.5e-168;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCGATCAAGAACACA 60
DB 4887 CACAAGTCACCAACAGCTGCAATCATACAGATGCAACAGCCGATCAAGAACACA 4946

QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGTCTTCAATCTCTCGAA 120
DB 4947 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGTCTTCAATCTCTCGAA 5006

QY 121 ATTACATCAGAAACCAACCATACATAGCTTCAACACACAGAGTCAAGTCAAACTG 180
DB 5007 ATTACATCAGAAACCAACCATACATAGCTTCAACACACAGAGTCAAGTCAAACTG 5066

QY 181 CAACCCACACAGTCAGATCTTAAACACACACACACACACACACACACACACACAC 240
DB 5067 CAATCCACACAGTCAGATCTTAAACACACACACACACACACACACACACACACAC 5126

QY 241 ACTCAAAACCAACGCCCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 300
DB 5127 ACCAATAACCAACGCCCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 5186

QY 301 GTGTTTAACTTTGACCTGAGCATATGAGGCAACCAATCAACCTCTGGCTATCTGC 360
DB 5187 GTGTTTAACTTTGACCTGAGCATATGAGGCAACCAATCAACCTCTGGCTATCTGC 5246

QY 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 420
DB 5247 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAAGCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 480
DB 5307 ACCTTCAAGCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCACGAGCCCAAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 5367 ACCACGAGCCCAAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTCAGCAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA 600
DB 5427 CTGCTCAGCAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA 5486

QY 601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660
DB 5487 TCAACCTCTCTCGAAGGCAATCCAGCCCTTCTCAAGTCTCTACACATCCGAGTACCCA 5546

QY 661 TCACACCTCTATCTCCACCCCAACCAACCAAGGCAAGTATTAATAAAAAA 714
DB 5547 TCACACCTCTATCTCCACCCCAACCAACCAAGGCAAGTATTAATAAAAAA 5600
```

```

RESULT 15
US-08-720-132-1
; Sequence 1, Application US/08720132
; Patent No. 6264957
; GENERAL INFORMATION:
; APPLICANT: Collins, Peter L.
; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY
; TITLE OF INVENTION: SYNCYTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
```

```

; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,132
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/007,083
; FILING DATE: 27-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 15280-250-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other
US-08-720-132-1
```

```

Query Match          91.6%; Score 654.8; DB 3; Length 15223;
Best Local Similarity 94.8%; Pred. No. 1.5e-168;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACTGCAATCATACAGATGCAACAGCCGATCAAGAACACA 60
DB 4887 CACAAGTCACCAACAGCTGCAATCATACAGATGCAACAGCCGATCAAGAACACA 4946

QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGTCTTCAATCTCTCGAA 120
DB 4947 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGTCTTCAATCTCTCGAA 5006

QY 121 ATTACATCAGAAACCAACCATACATAGCTTCAACACACAGAGTCAAGTCAAACTG 180
DB 5007 ATTACATCAGAAACCAACCATACATAGCTTCAACACACAGAGTCAAGTCAAACTG 5066

QY 181 CAACCCACACAGTCAGATCTTAAACACACACACACACACACACACACACACACAC 240
DB 5067 CAATCCACACAGTCAGATCTTAAACACACACACACACACACACACACACACACAC 5126

QY 241 ACTCAAAACCAACGCCCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 300
DB 5127 ACCAATAACCAACGCCCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 5186

QY 301 GTGTTTAACTTTGACCTGAGCATATGAGGCAACCAATCAACCTCTGGCTATCTGC 360
DB 5187 GTGTTTAACTTTGACCTGAGCATATGAGGCAACCAATCAACCTCTGGCTATCTGC 5246

QY 361 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 420
DB 5247 AAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 5306

QY 421 ACCTTCAAGCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 480
DB 5307 ACCTTCAAGCAACCAAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCA 5366

QY 481 ACCACGAGCCCAAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 5367 ACCACGAGCCCAAGAGGCAACCATCAACCAACCAACCAACCAACCAACCAACCA 5426

QY 541 CTGCTCAGCAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA 600
DB 5427 CTGCTCAGCAACCAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA 5486
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:50:52 ; Search time 244.018 Seconds  
(without alignments)  
7909.644 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715

Sequence: 1 cacaagaagtcactaacac.....gtagttattaaaaaaaaa 715

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.\*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	715	20	AXX08422
2	715	100.0	920	14	AAQ45686
3	715	100.0	920	20	AXX08421
4	659	92.2	935	8	AAV70784
5	659	92.2	935	19	AAV18736
6	657.4	91.9	935	13	AAQ29623
7	654.8	91.6	15210	20	AXX59703
8	654.8	91.6	15222	18	AAV78440

9	654.8	91.6	15223	18	AAV63430	Respiratory syncyt
10	654.8	91.6	15223	19	AAV17553	Respiratory syncyt
11	654.8	91.6	15223	21	AAV88743	Respiratory syncyt
12	649.6	90.9	918	22	AAV88494	Human RSV G-protei
13	646.2	90.4	897	24	AAV20145	Respiratory syncyt
14	645.8	90.4	897	19	AAV38298	Respiratory syncyt
15	645.8	90.4	897	19	AAV38298	Respiratory syncyt
16	300.4	42.0	15225	19	AAV17552	Respiratory syncyt
17	300.4	42.0	15225	19	AAV17552	Respiratory syncyt
18	295.6	41.3	15229	24	AAV18276	Human respiratory
19	295.6	41.3	15229	24	AAV18276	RSV isolate 18537
20	295.6	41.3	15229	20	AAV22910	Nucleotide sequenc
21	295.6	41.3	15229	20	AAV22910	DNA encoding the L
22	273.6	38.3	15218	19	AAV18275	RSV isolate 2B wil
23	273.6	38.3	15218	19	AAV18275	Nucleotide sequenc
24	273.6	38.3	15218	20	AAV22914	Nucleotide sequenc
25	273.6	38.3	15218	20	AAV22914	DNA encoding the L
26	273.6	38.3	15219	19	AAV18277	RSV vaccine 2B33F
27	273.6	38.3	15219	19	AAV18277	RSV vaccine 2B20L
28	273.6	38.3	15219	19	AAV18279	RSV revertant 2B33
29	273.6	38.3	15219	19	AAV18279	RSV revertant 2B20
30	273.6	38.3	15219	20	AAV22911	Nucleotide sequenc
31	273.6	38.3	15219	20	AAV22912	Nucleotide sequenc
32	273.6	38.3	15219	20	AAV22913	Nucleotide sequenc
33	273.6	38.3	15219	20	AAV22913	DNA encoding the L
34	273.6	38.3	15219	20	AAV22913	DNA encoding the L
35	273.6	38.3	15219	20	AAV22913	DNA encoding the L
36	212.8	29.8	1050	17	AAV1647	Encodes Streptococ
37	212.8	29.8	1050	22	AAV84711	Nucleotide sequenc
38	212.8	29.8	1356	22	AAV80153	Nucleotide sequenc
39	210.8	29.5	303	16	AAV03486	RSV subgroup A clo
40	210.8	29.5	303	17	AAV27066	RSV subgroup A prote
41	210.8	29.5	303	17	AAV27073	RSV subgroup A wi
42	210.8	29.5	303	17	AAV27073	Respiratory syncyt
43	210.8	29.5	303	17	AAV27073	RSV G protein anti
44	210.8	29.5	303	21	AAV5882	DNA encoding a G2N
45	210.8	29.5	303	22	AAV78459	Nucleotide sequenc

# ALIGNMENTS

RESULT 1  
AXX08422  
ID AXX08422 standard; cDNA; 715 BP.  
XX  
AC AXX08422;  
XX  
DT 28-JUN-1999 (first entry)  
XX  
DE G protein gene fragment of respiratory syncytial virus.  
XX  
KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
KW vaccine; immune response; immunogenicity; tPA; antibody;  
KW tissue plasminogen activator; ss.  
XX  
OS Respiratory syncytial virus (RSV).  
XX  
FH Key Location/Qualifiers  
FT CDS 1..702  
FT /tag= a  
FT /product= "Secreted G protein"

PN WO9904010-A1.

XX 28-JAN-1999.

XX 16-JUL-1998; 98WO-CA00697.

XX 18-JUL-1997; 97US-0896442.

XX (CONN-) CONNAUGHT LAB LTD.



PI Klein MH, Li X, Sambhara S;  
XX WPI: 1999-132254/11.  
DR P-PSDB; AAW96314.  
XX Immunogenic composition for generating antibodies against  
PT respiratory syncytial virus - comprises non-replicating vector  
PT containing the protein G sequence, useful in protective vaccines and  
PT to raise antibodies for diagnosis  
XX  
XX Claim 8; Figure 3; 67pp; English.  
XX  
XX The respiratory syncytial virus (RSV) G protein can be used in  
CC vaccines by inserting the G protein gene into a non-replicating  
CC vector. The G protein is placed under the control of alternative  
CC signal and expression sequences, for example the chimeric G protein  
CC produced may also comprise the signal peptide of tissue plasminogen  
CC activator (tPA). The recombinant vector may also comprise sequences  
CC upstream of the G protein gene which enhance the G proteins  
CC immunoprotective ability. The resulting immunogenic composition will  
CC generate antibodies directed against the RSV G protein when  
CC administered to a host organism. The composition is useful as a  
CC vaccine to immunise against RSV-associated disease, particularly  
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
CC by usual immunisation and cell fusion methods.  
XX  
XX Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;

Query Match 100.0%; Score 715; DB 20; Length 715;  
Best Local Similarity 100.0%; Pred. No. 1.4e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACACTGCAATCATACAGATGCAAGCCAGATCAAGACACA 60  
DB 1 CACAAAGTCACACTAACACTGCAATCATACAGATGCAAGCCAGATCAAGACACA 60  
QY 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120  
DB 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120  
QY 121 ATTACATCAAAACCCACCACTACTAGTTCACACACACAGAGTCAAGTCAAACTG 180  
DB 121 ATTACATCAAAACCCACCACTACTAGTTCACACACACAGAGTCAAGTCAAACTG 180  
QY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
DB 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
QY 241 ACTACAAACCAACCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
DB 241 ACTACAAACCAACCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
QY 301 GTGTTTAACTTTTACCTGACATATGAGCAATCAGCAATCAGCTGCTGGGTATCTGC 360  
DB 301 GTGTTTAACTTTTACCTGACATATGAGCAATCAGCAATCAGCTGCTGGGTATCTGC 360  
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
DB 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
DB 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
QY 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
DB 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
DB 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
QY 601 TCAACCTCTCTCGAAGGCAATCTTAAGCCCTTTCTCAAGTCTTCCAAACATCCGAGCACCA 660

DB 601 TCAACCTCTCTCGAAGGCAATCTTAAGCCCTTTCTCAAGTCTTCCAAACATCCGAGCACCA 660  
QY 661 TCACAAACCTCTCTCTCCACCCACCAACCAACCAACCAACCAACCAACCAACCA 715  
DB 661 TCACAAACCTCTCTCTCCACCCACCAACCAACCAACCAACCAACCAACCAACCA 715  
RESULT 2  
AAQ45686  
ID AAQ45686 standard; DNA; 920 BP.  
XX  
XX AAQ45686;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 13-JAN-1994 (first entry)  
XX  
XX Respiratory syncytial virus (RSV) G gene.  
XX PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine;  
XX db.  
XX Respiratory syncytial virus.  
XX OS  
XX Key Location/Qualifiers  
XX CDS 8..901  
XX FT /\*tag= a  
XX FT /product= RSV G protein  
XX FT 89..176  
XX FT /\*tag= b  
XX FT /label= Transmembrane anchor domain  
XX  
XX WO9314207-AL.  
XX  
XX 22-JUL-1993.  
XX  
XX 05-JAN-1993; 93WO-CA00001.  
XX  
XX 06-JAN-1992; 92GB-0000117.  
XX  
XX (CONN-) CONNAUGHT LAB LTD.  
XX  
XX Ewasysbyn ME, Klein MH;  
XX  
XX WPI: 1993-243222/30.  
XX P-PSDB; AAR39286.  
XX  
XX Multimeric hybrid genes and their chimeric proteins - are  
XX PT vaccines against multiple pathogenic infections e.g.  
XX PT para-influenza virus and respiratory syncytial virus  
XX  
XX Claim 11; Figure 7A-7D; 80pp; English.  
XX  
XX A novel multimeric hybrid gene is used as a vaccine. The gene  
XX consists of two gene sequences which are linked and encode antigenic  
XX regions, these two sequences being derived from two different  
XX pathogens (para-influenza virus (PIV) and respiratory syncytial virus  
XX (RSV)). The gene sequences that are particularly used are those  
XX which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
XX G proteins (AAQ45685, AAQ45686).  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 715; DB 14; Length 920;  
Best Local Similarity 100.0%; Pred. No. 1.5e-162;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAACACTGCAATCATACAGATGCAAGCCAGATCAAGACACA 60  
DB 206 CACAAAGTCACACTAACACTGCAATCATACAGATGCAAGCCAGATCAAGACACA 265  
QY 61 ACCCCAAATACCTCACTCAGGATCTCTGGAATCAGCTTCTCCAATCTGTCTGAA 120

```

Db 266 ACCCCAAATACATCTCACTCAGGATCTCAGCTTGGAAATCAGCTTCTCCAATCTGTCTGAA 325
Qy 121 ATTATCATCAAAACCAACCACTACTAGTCTCAACACACACAGGAGTCAAGTCAAACTGT 180
Db 326 ATTACATCAAAACCAACCACTACTAGTCTCAACACACACAGGAGTCAAGTCAAACTGT 385
Qy 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
Db 386 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 445
Qy 241 ACTACAAACACAGGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
Db 446 ACTACAAACACAGGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 505
Qy 301 GTGTTTAACTTGTACCTGAGCATATGCGACCAATCCACCTGCTGGCTATCTGTC 360
Db 506 GTGTTTAACTTGTACCTGAGCATATGCGACCAATCCACCTGCTGGCTATCTGTC 565
Qy 361 AAAAGAATACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 420
Db 566 AAAAGAATACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 625
Qy 421 ACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480
Db 626 ACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 685
Qy 481 ACCACCAAGCCACAGAGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAAC 540
Db 586 ACCACCAAGCCACAGAGAGCCACCATCAACACCAACCAACCAACCAACCAACCAACCAAC 745
Qy 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
Db 746 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805
Qy 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 660
Db 806 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 865
Qy 661 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 715
Db 866 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 920

```

## RESULT 3

```

AAx08421
ID AAx08421 standard; cDNA; 920 BP.
AC AAx08421;
XX
XX
DT 28-JUN-1999 (first entry)
XX
DE Membrane bound G protein gene of respiratory syncytial virus.
XX
KW G protein; respiratory syncytial virus; RSV; recombinant vector;
KW vaccine; immune response; immunogenicity; tPA; antibody;
KW tissue plasminogen activator; ss.
XX
OS Respiratory syncytial virus (RSV).
XX
XX Key Location/Qualifiers
XX CDS 8..904
XX FT /*tag= a
XX FT /product= "Membrane bound G protein"
XX
XX W09904010-A1.
XX
XX 28-JAN-1999.
XX
XX 16-JUL-1998; 98WO-CA00697.
XX
XX 18-JUL-1997; 97US-0896442.
XX

```

PA

(CONN-) CONNAUGHT LAB LTD.

Klein MH, Li X, Sambhara S;

WPI; 1999-132254/11.

P-PSDB; AAW96313.

Immunogenic composition for generating antibodies against

respiratory syncytial virus - comprises non-replicating vector containing the protein G sequence, useful in protective vaccines and to raise antibodies for diagnosis

Claim 3; Figure 2; 67pp; English.

CC The respiratory syncytial virus (RSV) G protein can be used in vaccines by inserting the G protein gene into a non-replicating vector. The G protein is placed under the control of alternative signal and expression sequences, for example the chimeric G protein produced may also comprise the signal peptide of tissue plasminogen activator (tPA). The recombinant vector may also comprise sequences upstream of the G protein gene which enhance the G proteins immunoprotective ability. The resulting immunogenic composition will generate antibodies directed against the RSV G protein when administered to a host organism. The composition is useful as a vaccine to immunise against RSV-associated disease, particularly resulting in a balanced Th1/Th2 immune response and for raising Ab, by usual immunisation and cell fusion methods.

Sequence 920 BP; 380 A; 230 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 715; DB 20; Length 920;

Best Local Similarity 100.0%; Pred. No. 1.5e-162;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 CACAAGTCACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60
Db 206 CACAAGTCACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 265
Qy 61 ACCCAACATCTCACTCAGGATCTCAGCTTGGAAATCAGTCTTCCAATCTCTGAA 120
Db 266 ACCCAACATCTCACTCAGGATCTCAGCTTGGAAATCAGTCTTCCAATCTCTGAA 325
Qy 121 ATTATCATCAAAACCAACCACTACTAGTCTCAACACACAGGAGTCAAGTCAAACTG 180
Db 326 ATTATCATCAAAACCAACCACTACTAGTCTCAACACACAGGAGTCAAGTCAAACTG 385
Qy 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
Db 386 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 445
Qy 241 ACTACAAACCAACCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 446 ACTACAAACCAACCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
Qy 301 GTGTTTAACTTGTACCTGAGCATATGCGACCAATCCACCTGCTGGCTATCTGTC 360
Db 506 GTGTTTAACTTGTACCTGAGCATATGCGACCAATCCACCTGCTGGCTATCTGTC 565
Qy 361 AAAAGAATACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 420
Db 566 AAAAGAATACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 625
Qy 421 ACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480
Db 626 ACCTTCAAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 685
Qy 481 ACCACCAAGCCACAGAGGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAAC 540
Db 686 ACCACCAAGCCACAGAGGCAACCATCAACACCAACCAACCAACCAACCAACCAACCAAC 745
Qy 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
Db 746 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 805

```

```

QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660
Db 806 TCAACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 865
QY 661 TCACAAACCTCTCATCTCCACCCACACACACACGCGCAGTAGTTATTAAAAA 715
Db 866 TCACAAACCTCTCATCTCCACCCACACACACACGCGCAGTAGTTATTAAAAA 920

RESULT 4
AAV18736
ID AAV18736 standard; cDNA; 935 BP.
XX
AC AAV18736;
XX
XX 25-MAR-2003 (updated)
DT 05-APR-1991 (first entry)
XX
XX Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE G protein.
DE
XX
XX Vaccine; ss.
XX
XX Human respiratory syncytial virus (HRSV).
XX
FH Key Location/Qualifiers
FT CDS 16..913
FT /*tag= a
XX
XX WO8704185-A.
XX
XX 16-JUL-1987.
XX
XX 23-DEC-1986; 86WO-US02756.
XX
XX 14-JAN-1986; 86US-0818740.
XX
XX (UNYNC-) UNIV NORTH CAROLINA.
PA (WERTZ/) WERTZ G W.
XX
XX WPI; 1987-206300/29.
DR P-PSDB; AAP70845.
XX
XX Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
XX
XX Disclosure; Chart 13; 57pp; English.
XX
XX A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;
SQ

Query Match 92.2%; Score 659; DB 8; Length 935;
Best Local Similarity 95.1%; Pred. No. 4.5e-149;
Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACAATGCAATCATACAGATGCAACAAGCCAGATCAAGAACA 60
Db 214 CACAAAGTCACACCAACAATGCAATCATACAGATGCAACAAGCCAGATCAAGAACA 273
QY 61 ACCCCACATCTCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGT 120
Db 274 ACCCCACATCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGT 333
QY 121 ATTACATCAACAACCAACCACTACTAGTCTTCAACACACGAGTCAAGTCAACCTG 180

```

```

Db 334 ATTACATCAACAATCACCACCATCTAGCTTCAACAACACGAGGATCAAGTCAACCCCTG 393
QY 181 CAAACCAACAGTCAAGTCTAAGTCTAAGTCTAAGTCTAAGTCTAAGTCTAAGTCT 240
Db 394 CAATCCACACAGTCAAGTCAAGTCTAAGTCTAAGTCTAAGTCTAAGTCTAAGTCT 453
QY 241 ACTACAAACAGGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 454 ACCACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513
QY 301 GTGTTTAACTTTGTATCCCTGACGATATGACGACCAACCAACCAACCAACCAACCA 360
Db 514 GTGTTCAACTTTGTATCCCTGACGATATGACGACCAACCAACCAACCAACCAACCA 573
QY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 574 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
Db 634 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 693
QY 481 ACCACCAAGCCACAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 540
Db 694 ACCACCAAGCCACAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 753
QY 541 CTGCTCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
Db 754 CTACTCACCTCCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 813
QY 601 TCACAACTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 660
Db 814 TCACAACTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGCACCA 873
QY 661 TCACAACTCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
Db 874 TCACAACTCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 928

RESULT 5
AAV18736
ID AAV18736 standard; cDNA; 935 BP.
XX
AC AAV18736;
XX
XX 11-JUN-1998 (first entry)
DT
DE HRSV glycoprotein G cDNA.
XX
XX HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine; ss.
XX
XX Human respiratory syncytial virus.
XX
XX Key Location/Qualifiers
XX CDS 16..912
XX /*tag= a
XX /product= glycoprotein_G
XX
XX US5716823-A.
XX
XX 10-FEB-1998.
XX
XX 12-MAY-1997; 97US-0854783.
XX
XX 13-JUL-1988; 88US-0218737.
XX
XX 14-JAN-1986; 86US-0818740.
XX
XX 23-DEC-1986; 86WO-US02756.
XX
XX 11-JUN-1992; 92US-0897171.
XX
XX 12-MAY-1997; 97US-0854783.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Collins PL, Wertz GW;
PI

```

```

XX WP1; 1998-144802/13.
DR P-PSDB; AAW47605.
XX
XX Production of human respiratory syncytial virus glyco-protein P or G
PT -by culturing eukaryotic host cells transfected with corresponding
PT DNA
XX
XX Example 1; Columns 27-28; 17pp; English.
XX
XX The present sequence was used in the development of a novel method
CC for the production of human respiratory syncytial virus (HRSV)
CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
CC culturing eukaryotic host cells transfected with an isolated DNA
CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare
CC vaccines against HRSV.
XX
XX Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;
SQ
Query Match 92.2%; Score 659; DB 19; Length 935;
Best Local Similarity 95.1%; Pred. No. 4.5e-149;
Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAAACAACTGCAATCATATACAGATGCAACAGCCAGATCAAGAACACA 60
DB 214 CACAAAGTCACACCACTGCAATCATATACAGATGCAACAGCCAGATCAAGAACACA 273
QY 61 ACCCAACATPACTCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGCTGAA 120
DB 274 ACCCAACATPACTCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGCTGAA 333
QY 121 ATTACATCACAACACCACTACTAGTCTTCAACACCAACCAAGTCAAGTCAACCTG 180
DB 334 ATTACATCACAACATCACCACCACTACTAGTCTTCAACACCAACCAAGTCAAGTCAACCTG 393
QY 181 CAACCCACACAGCTCAGACTAATAAACAACAACCAACCAACCAACCAACCAACCAACCA 240
DB 394 CAATCCACACAGCTCAGACTAATAAACAACAACCAACCAACCAACCAACCAACCAACCA 453
QY 241 ACTACAAAACAGCCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 454 ACCACAAAACAGCCAAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513
QY 301 GTGTTTAACTTGTACCTCAGCATATGCAAGCAACCAACCAACCAACCAACCAACCAACCA 360
DB 514 GTGTTTAACTTGTACCTCAGCATATGCAAGCAACCAACCAACCAACCAACCAACCAACCA 573
QY 361 AAAGAATATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 574 AAAGAATATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 634 ACCCTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693
QY 481 ACCACCAAGCCCAACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 694 ACCACCAAGCCCAACCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
DB 754 CTACTCACTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 813
QY 601 TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCCAACCAACCAACCAACCA 660
DB 814 TCAACTTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCTCAACCAACCAACCAACCA 873
QY 661 TCACCAACCTCATCTCCACCCCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 874 TCACCAACCTCATCTCCACCCCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 928

```

RESULT 6

```

AAQ29623
ID AAQ29623 standard; DNA; 935 BP.
XX
XX AC AAQ29623;
XX
XX 25-MAR-2003 (updated)
DT 03-MAR-1993 (first entry)
XX
XX HRSV glycoprotein G (gpG).
XX
XX Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
XX major capsid protein; N; ss.
XX
XX Human respiratory syncytial virus strain A2.
XX
XX Key Location/Qualifiers
FT CDS 16..912
FT /tag= a
FT /label= G_protein
FT misc_feature 16..22
FT /tag= b
FT /note= "Oligonucleotide used to probe for
FT full length cDNA"
FT misc_feature 898..912
FT /tag= c
FT /note= "Oligonucleotide used to specifically
FT prime the reverse transcription reaction
FT for making the first strand of the cDNA"
XX
XX US5149650-A.
XX
XX 22-SEP-1992.
XX
XX 13-JUL-1988; 88US-0218737.
XX
XX 14-JAN-1986; 86US-0818740.
XX 13-JUL-1988; 88US-0218737.
XX
XX (UYN-C) UNIV NORTH CAROLINA.
XX
XX Collins PL, Wertz GW;
XX
XX WP1; 1992-340247/41.
XX P-PSDB; AAR25302.
XX
XX Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
XX
XX Disclosure; Page 18; 21pp; English.
XX
XX The sequences of mRNA encoding HRSV structural proteins are given in
CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;
SQ
Query Match 91.9%; Score 557.4; DB 13; Length 935;
Best Local Similarity 95.0%; Pred. No. 1.1e-148;
Matches 679; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAAACAACTGCAATCATATACAGATGCAACAGCCAGATCAAGAACACA 60
DB 214 CACAAAGTCACACCACTGCAATCATATACAGATGCAACAGCCAGATCAAGAACACA 273
QY 61 ACCCAACATPACTCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGCTGAA 120
DB 274 ACCCAACATPACTCTCACTCAGGATCCTCAGCTTGGAAATCAGTTCTCCAAATCTGCTGAA 333
QY 121 ATTACATCACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180

```



AC AAT78440;  
 XX 17-MAR-1998 (first entry)  
 XX Human respiratory syncytial virus strain A2.  
 XX Antisense oligonucleotide; respiratory syncytial virus;  
 KW RSV; treatment; infection; inhibition; strain A2; ss.  
 XX Human respiratory syncytial virus.  
 OS WO9729757-A1.  
 XX 21-AUG-1997.  
 XX 14-FEB-1997; 97WO-US02531.  
 XX 15-FEB-1996; 96US-0011725.  
 XX (CLEV-) CLEVELAND CLINIC FOUND.  
 PA (USSH ) US NAT INST OF HEALTH.  
 XX Cirino NM, Li G, Silverman RH, Torrence PF, Xiao W;  
 XX WPI; 1997-424748/39.  
 DR Polynucleotide containing sequence anti-sense to region of RSV -  
 XX connected via a linker to an activator of RNaseL, used to treat RSV  
 XX infections  
 XX Disclosure; Pages 47-51; 89pp; English.  
 XX The present sequence was used in the preparation of a novel  
 CC polynucleotide, comprising an antisense oligonucleotide, with a  
 CC hydroxy group at one end, that is complementary to 15-20 bases of  
 CC the anti-genomic RNA strand of a respiratory syncytial virus (RSV),  
 CC a linker attached to the OH-end of the antisense oligonucleotide and  
 CC an oligonucleotide activator of RNaseL attached to the linker. The  
 CC polynucleotide can be used to treat RSV infections, which can also  
 CC be treated by administration of the antisense oligonucleotide, so as  
 CC to form a complex with activated RNase L in vivo. The  
 CC polynucleotide can be transported across the cell membranes without  
 CC carriers or permeability agents, and once introduced destroys  
 CC antisense target RNA. It also inhibits RSV infection in vitro in a  
 CC superior manner to the conventional drug, ribavirin.  
 XX  
 SQ Sequence 15222 BP; 5923 A; 2706 C; 2356 G; 4237 T; 0 other;  
 Query Match 91.6%; Score 654.8; DB 18; Length 15222;  
 Best Local Similarity 94.8%; Pred. No. 9.4e-148;  
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
 QY 1 CACAAAGTCACACTAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 60  
 DB 4886 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACACA 4945  
 QY 61 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGAATCAGCTTCTCAATCTGCTGAA 120  
 DB 4946 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGAATCAGCTTCTCAATCTGCTGAA 5005  
 QY 121 ATTACATCACAACCCACCACCATCTAGCTTCAACACACAGGAGTCAAGTCAAACTG 180  
 DB 5006 ATTACATCACAACATCACCACCATCTAGCTTCAACACACAGGAGTCAAGTCAACCTG 5065  
 QY 181 CAACCCACACAGTCAAGACTAAACACACACACACACACACACACACACACACACACAC 240  
 DB 5066 CAATCCACACAGTCAAGACCAACACACACACACACACACACACACACACACACAC 5125  
 QY 241 ACTACAAAACAGCCCAAC 300  
 DB 5126 ACCACAAAACAGCCCAAC 5185  
 QY 301 GTGTTTAACTTTGTACCTGCAGCATATGCGACACAACTCAACCTGCTGGCTATCTGC 360

Db 5186 GTGTTCACTTTGTACCTGCAGCATATGCGACAACTCAACCTGCTGGCTATCTGC 5245  
 QY 361 AAAAGAATACCAACCAAAAAACAGGAAAGAAAACCCACCAACAGCCTTACAAAAAACCA 420  
 Db 5246 AAAAGAATACCAACCAAAAAACAGGAAAGAAAACCCACCAACAGCCTTACAAAAAACCA 5305  
 QY 421 ACCTTCAGACACCAACCAAAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAACCTCA 480  
 Db 5306 ACCCTCAAGACACCAACCAAAAAAGATCTCAACCTCAACCTCAACCTCAACCTCAACCT 5365  
 QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACACACACACACACACACACACACACAC 540  
 Db 5366 ACCACCAAGCCACAGAGAGCCCAACCATCAACACACACACACACACACACACACACAC 5425  
 QY 541 CTGCTCACCAACACACACACAGGAAATCCAAAATCTCAAACTCAAACTCAAACTCAAACT 600  
 Db 5426 CTACTCACCTCCACACACACAGGAAATCCAAAATCCAAAATCCAAAATCCAAAATCC 5485  
 QY 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGCTCTCAAGTCTCCACCAACATCCGACCCA 660  
 Db 5486 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGCTCTCAAGTCTCTACACATCCGAGTACCA 5545  
 QY 661 TCACACCTCTCATCTCCACCCACACACACACACACACACACACACACACACACACAC 714  
 Db 5546 TCACACCTCTCATCTCCACCCACACACACACACACACACACACACACACACACACATA 5599  
 RESULT 9  
 AAT63430  
 ID AAT63430 standard; DNA, 15223 BP.  
 XX  
 AC AAT63430;  
 XX  
 DT 02-JUL-1997 (first entry)  
 XX  
 DE Respiratory syncytial virus anti-genome.  
 XX  
 KW RSV; vaccine; gene therapy; upper respiratory tract disease; ss.  
 XX  
 OS Human respiratory syncytial virus strain A2.  
 XX  
 PN WO9712032-A1.  
 XX  
 PD 03-APR-1997.  
 XX  
 PF 27-SEP-1996; 96WO-US15524.  
 XX  
 PR 27-SEP-1995; 95US-0007083.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Collins PL;  
 XX  
 WPI; 1997-212893/19.  
 XX  
 XX Infectious respiratory syncytial virus particles - useful for  
 treatment of RSV or gene therapy of upper respiratory tract diseases  
 Claim 46; Page 43-51; 66pp; English.  
 XX  
 CC A human respiratory syncytial virus (RSV) anti-genome sequence  
 (AAT63430) is the 5' to 3' positive-sense sequence of RSV; the genome  
 itself is negative-sense. It was synthesised in segments by RT-PCR  
 using intracellular RSV mRNA or genomic rRNA isolated from purified  
 viruses as template. Restriction site markers were introduced by  
 incorporating the changes into the primers used for RT-PCR. The  
 recombinant sequence can be expressed with a nucleocapsid protein,  
 a nucleocapsid phosphoprotein, a large polymerase protein and an  
 RNA elongation factor to produce isolated infectious RSV particles  
 useful for generating vaccines against RSV. Recombinant RSV genome  
 or antigenome can also be used as a vector for gene therapy of the  
 upper respiratory tract.

```

XX SQ Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;
Query Match      91.6%; Score 654.8; DB 18; Length 15223;
Best Local Similarity 94.8%; Pred. No. 9.4e-148;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACAAGTCGAATCATACAGATGCAACAGCCAGATCAAGAACACA 60
Db 4887 CACAAAGTCACAAACCAACCAAGTCGAATCATACAGATGCAACAGCCAGATCAAGAACACA 4946
QY 61 ACCCAACATACCTCACTCAGATCCTCAGCTTGGAAATCAGCTTCTCAATCTCTCTGAA 120
Db 4947 ACCCAACATACCTCACTCAGATCCTCAGCTTGGAAATCAGCTTCTCAATCTCTCTGAA 5006
QY 121 ATTACATCAAAACACCAACATCTAGCTTCAACAAACACAGAGTCAAGTCAAAACCTG 180
Db 5007 ATTACATCAAAATCACCACTACTAGCTTCAACAAACACAGAGTCAAGTCAAAACCTG 5066
QY 181 CAACCCACACAGTCAGATCAAAACACAAACACAAACCAACCAACCAACCAAGCC 240
Db 5067 CAATCCCAACAGTCAGATCAAAACCAAAACCAACCAACCAACCAACCAAGCC 5126
QY 241 ACTACAAACCAACGCCCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5127 ACCAACAACCAACGCCCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186
QY 301 GTGTTTAACTTTGACCTCGCATATGCAAGCAACATCCAACTGCTGGCTATCTGC 360
Db 5187 GTGTTTAACTTTGACCTCGCATATGCAAGCAACATCCAACTGCTGGCTATCTGC 5246
QY 361 AARAGATACCAACAAACCAACCAAGGAAGAAACCAACCAACCAACCAACCAACCAACCA 420
Db 5247 AAAAGATACCAACAAACCAACCAAGGAAGAAACCAACCAACCAACCAACCAACCAACCA 5306
QY 421 ACCTTCAAGCAACCAACCAAGATCTCAAACTTCAAACTTCAAACTTCAAACTTCAAACT 480
Db 5307 ACCTTCAAGCAACCAACCAAGATCTCAAACTTCAAACTTCAAACTTCAAACTTCAAACT 5366
QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 540
Db 5367 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCAAC 5426
QY 541 CTGCTCAGCAACACACACAGGAATTCAAAATCTCAAGTCAAAATGGAACCTTCCAC 600
Db 5427 CTACTCACTTCAACACACAGGAATTCAGAATCTCAAGTCAAAATGGAACCTTCCAC 5486
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCGAGCACCA 660
Db 5487 TCAACTTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCTCAAGTCTCTCAACATCGAGTACCA 5546
QY 661 TCACACCTCTATCTCCACCAACACACAGCCAGCTAGTTATTAAAAA 714
Db 5547 TCACACCTTCTATCTCCACCAACACACAGCCAGCTAGTTATTAAAAACATA 5600

```

RESULT 10

AAV17553

ID AAV17553 standard; cDNA; 15223 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

```

XX PF 15-JUL-1997; 97WO-US12269.
XX PR 23-MAY-1997; 97US-0047634.
XX PR 15-JUL-1996; 96US-0021773.
XX PR 09-MAY-1997; 97US-0046141.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN,
XX Whitehead SS;
XX WPI; 1998-110579/10.
XX Attenuated respiratory syncytial virus vaccines - useful to protect
XX individuals against RSV infection
XX Example 7; Page 188-195; 238pp; English.
XX This is the 5'-3' positive sequence nucleotide sequence of
XX respiratory syncytial virus (RSV) D46. The genome is
XX negative-sense; the complete nucleotide sequence of the
XX wild-type B-1 virus has also been determined (see AAV17552).
XX A novel infectious recombinant RSV comprises a RSV genome or
XX antigenome, a major nucleocapsid (N) protein, a nucleocapsid
XX phosphoprotein (P), a large polymerase protein (L), and a RNA
XX polymerase elongation factor, where the recombinant RSV has at
XX least two attenuating mutations, one of the mutations specifying a
XX temperature-sensitive (ts) substitution at amino acid Phe521.
XX Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts
XX nucleotide substitution in the gene-start sequence of gene M2.
XX Also claimed are: (i) an isolated infectious RSV particle which
XX comprises a recombinant RSV (antigenome, N, P, and L proteins, a
XX RNA polymerase elongation factor, where the (anti)genome is modified;
XX (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or
XX a cis-acting regulatory sequence; and (ii) by a termination codon
XX introduced within a selected gene, or by a change in sequence,
XX position or presence of a GS or GE transcription signal relative to
XX the selected gene; (2) an expression vector; and (3) an RSV strain
XX selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR
XX 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR
XX 2452), cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or
XX B-1 cp52/285 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated
XX attenuated recombinant RSV and RSV particles are used in a vaccine
XX to stimulate the immune system of an individual to induce
XX protection against RSV. The expression vector of (2) is used for
XX the production of infectious attenuated RSV particles.
XX SQ Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

```

```

Query Match      91.6%; Score 654.8; DB 19; Length 15223;
Best Local Similarity 94.8%; Pred. No. 9.4e-148;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

QY 1 CACAAAGTCACACTAACAAGTCGAATCATACAGATGCAACAGCCAGATCAAGAACACA 60
Db 4887 CACAAAGTCACAAACCAACCAAGTCGAATCATACAGATGCAACAGCCAGATCAAGAACACA 4946
QY 61 ACCCAACATACCTCACTCAGATCCTCAGCTTGGAAATCAGCTTCTCAATCTCTCTGAA 120
Db 4947 ACCCAACATACCTCACTCAGATCCTCAGCTTGGAAATCAGCTTCTCAATCTCTCTGAA 5006
QY 121 ATTACATCAAAACACCAACATCTAGCTTCAACAAACACAGAGTCAAGTCAAAACCTG 180
Db 5007 ATTACATCAAAATCACCACTACTAGCTTCAACAAACACAGAGTCAAGTCAAAACCTG 5066
QY 181 CAACCCACACAGTCAGATCAAAACACAAACACAAACCAACCAACCAACCAAGCC 240
Db 5067 CAATCCCAACAGTCAGATCAAAACCAAAACCAACCAACCAACCAACCAAGCC 5126
QY 241 ACTACAAACCAACGCCCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5127 ACCAACAACCAACGCCCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5186

```

20-JUL-1998 (first entry)

Respiratory syncytial virus antigenome.

RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.

Human respiratory syncytial virus D46.

WO9802530-A1.

22-JAN-1998.







Db 5247 AAAAGAAATACCAACCAAAAAACGAGAAAGAAAAACCTACCAAGCCCAAAAAAACCA 5306  
Qy 421 ACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAAGTACCC 480  
Db 5307 ACCCTCAAGACAAACCAAAAAAGATCCCAAACTCAAAACCACTAAACCAAGGAAGTACCC 5366  
Qy 481 ACCACCAAGCCCAAGAGAGGCAACCACTCAACACCAACCAAAAAAATCAACAATCAACTACA 540  
Db 5367 ACCACCAAGCCCAAGAGAGGCAACCACTCAACACCAACCAAAAAAATCAACAATCAACTACA 5426  
Qy 541 CTGCTCACCACCAACACCAAGGAAATCCAAACTCAAGCTCAAGTCAATGGAACCTTCCAC 600  
Db 5427 CTACTCACTCCACACCAAGGAAATCCAAACTCAAGTCAATGGAACCTTCCAC 5486  
Qy 601 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACCAACATCCGAGCACCCA 660  
Db 5487 TCAACTCTCTCGAAGGCAATCCAAAGCCCTTCTCAAGTCTCTACAAACATCCGAGTACCCA 5546  
Qy 661 TCACAACTCTATCTCCACCAACCAACACCAAGGAGTATTAATAAAAAAAA 714  
Db 5547 TCACAACTCTATCTCCACCAACCAACACCAAGGAGTATTAATAAAAAACA 5600

RESULT 12

AAC88494  
ID AAC88494 standard; DNA; 918 BP.  
XX  
AC AAC88494;  
DT 12-MAR-2001 (first entry)  
XX  
DE Human RSV G-protein gene.  
XX  
KW Chimeric; respiratory syncytial virus; RSV; immunize; ds.  
XX  
OS Unidentified.  
XX  
FN WO200068392-A1.  
XX  
PD 16-NOV-2000.  
XX  
PF 10-MAY-2000; 2000WO-US12582.  
XX  
PR 11-MAY-1999; 99US-0133536.  
XX  
PA (UNITI ) UNIV ILLINOIS FOUND.  
XX  
PI Buetow DE, Korban SS, Sandhu J, Krasnyanski SF;  
XX  
DR WPI; 2001-122707/13.  
XX  
PT Chimeric nucleic acid construct for immunizing animals and humans  
PT against respiratory syncytial virus (RSV), comprises a sequence adapted  
PT for expression in plants and a RSV protein or peptide coding sequence  
PT

Disclosure; Fig 10; 67pp; English.

XX  
CC The present invention relates to a chimeric nucleic acid construct  
CC comprising: a nucleotide sequence adapted for protein expression in  
CC plants; and a respiratory syncytial virus (RSV) coding sequence  
CC encoding an RSV protein or an antigenic protein or peptide of RSV.  
CC The construct can be used to immunize animals and humans against  
CC respiratory syncytial virus. The use of transgenic plants to  
CC generate the antigen allows the production of greater amounts of  
CC antigen.  
XX  
SQ Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;

Query Match 90.9%; Score 649.6; DB 22; Length 918;  
Best Local Similarity 95.2%; Pred. No. 8.1e-147;  
Matches 670; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1 CACAAAGTCACTAAACAACCTGCATCATACAAGATGCAACAAGCCAGATCAAGAACA 60  
Db 214 CACAAAGTCACTAAACAACCTGCATCATACAAGATGCAACAAGCCAGATCAAGAACA 273  
Qy 61 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGATTCAGCTTCTCCATCTCTGAA 120  
Db 274 ACCCCACATACCTCAGTCAGGATCCCTCAGCTTGGATTCAGCTTCTCCATCTCTGAA 333  
Qy 121 ATTACATCAAAAACACCACTACTAGCTTCAACAACACAGGAGTCAAGTCAAACTG 180  
Db 334 ATTACATCAAAAACACCACTACTAGCTTCAACAACACAGGAGTCAAGTCAAACTG 393  
Qy 181 CAACCCACAGCTCAAGATTAATAACAACAACAACCCAAAACAACCCAGAGGCC 240  
Db 394 CAATCCCAACAGTCAAGACCAAAAACAACAACAACCACTCAACCAACCCAGAGGCC 453  
Qy 241 ACTCAAAAACAACCCAAAACAACAACAACCCAAAACAACCCAAATATGATTTCACTTGA 300  
Db 454 ACCAATAAACGCAACCAACCAACCAACCAACCAACCAACCAATATGATTTCACTTGA 513  
Qy 301 GTGTTTAACTTTGTACCTCGAGCATATGCAGCAACAATCCAACTGCTGGGCTATCTGC 360  
Db 514 GTGTTTAACTTTGTACCTCGAGCATATGCAGCAACAATCCAACTGCTGGGCTATCTGC 573  
Qy 361 AAAAGATATCAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
Db 574 AAAAGATATCAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633  
Qy 421 ACCTTCAAGACAAACCAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACT 480  
Db 634 ACCTTCAAGACAAACCAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACT 693  
Qy 481 ACCACCAAGCCCAAGAGAGCCCAACCATCAACACCAACCAACCAACCAACCAACCA 540  
Db 694 ACCACCAAGCCCAAGAGAGCCCAACCATCAACACCAACCAACCAACCAACCAACCA 753  
Qy 541 CTGCTCACCACCAACCAACCAAGGAAATCCAAACTCAAGTCAAACTCAAACTCAAACT 600  
Db 754 CTACTCACCACCAACCAACCAAGGAAATCCAAACTCAAGTCAAACTCAAACTCAAACT 813  
Qy 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 660  
Db 814 TCACCTCTCCGAGGCAATCTAAGCCCTTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 873  
Qy 661 TCACACCTCTCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 704  
Db 874 TCACACCTCTCTCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 917

RESULT 13

AAS20145  
ID AAS20145 standard; DNA; 897 BP.  
XX  
AC AAS20145;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Respiratory syncytial virus G protein DNA.

XX  
KW RSV; ds; G protein; heavily glycosylated protein; antianemic; antiviral;  
KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
KW antiviral chemotherapeutic compound; humoral response;  
KW cellular immune response; hPIV; paediatric respiratory disease;  
KW globulin gene transfer; sickle cell disease; beta-thalassaemia;  
KW human immunodeficiency virus infection; HIV.  
OS Human respiratory syncytial virus.  
XX  
FH Key Location/Qualifiers  
XX CDS 1..897  
XX FT /\*tag= a  
XX FT /product= "G protein"

XX WO200192548-A2.  
 XX 06-DEC-2001.  
 XX 22-MAY-2001; 2001WO-US16610.  
 XX 01-JUN-2000; 2000US-208701P.  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Portner A, Takimoto T;  
 XX WPI; 2002-130534/17.  
 XX P-PSDB; AAU74676.  
 XX Recombinant Sendai virus useful in vaccines to protect infection by  
 XX paramyxoviruses, comprises exogenous nucleic acid encoding  
 XX paramyxovirus protein or its antigenic fragment -  
 XX Disclosure; Page 47; 57pp; English.  
 XX The invention relates to a recombinant Sendai virus comprising an  
 XX exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 XX antigenic fragment. The virus may be administered in combination  
 XX with an antiviral chemotherapeutic compound. Two or more viruses  
 XX expressing different PMV proteins may be co-administered. Compositions  
 XX comprising the virus are useful for eliciting a humoral and/or  
 XX cellular immune response to a PMV in a mammal, particularly a human.  
 XX Further a recombinant Sendai virus comprising an exogenous nucleic acid  
 XX encoding a second PMV protein is also administered and priming and/or  
 XX boosting humoral or cellular immune response comprises administering  
 XX one or more of a recombinant or isolated PMV protein or its antigenic  
 XX fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
 XX vector encoding a PMV protein. The recombinant virus is useful as an  
 XX effective vaccine against hPIV or RSV (the major causes of paediatric  
 XX respiratory disease) and also to express any gene of  
 XX interest in target cells, providing a positive medical impact on  
 XX impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 XX into stem cells effects a cure for sickle cell disease or beta-  
 XX thalassaemia. The recombinant virus may also prove effective in  
 XX conferring immunity to human immunodeficiency virus (HIV) infection.  
 XX The Sendai virus replicates at level that is high enough to  
 XX induce sufficient immunity, but does not cause any harm to human  
 XX recipient. The present sequence encodes a respiratory syncytial  
 XX virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 XX suitable for expression by the recombinant virus of the invention.  
 XX SQ Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other;  
 Query Match 90.4%; Score 646.2; DB 24; Length 897;  
 Best Local Similarity 95.3%; Pred. No. 5.3e-146;  
 Matches 666; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
 QY 1 CCAAAAGTCACATCAACATCGCAATCATACAGATGCAAGCCAGATCAAGAACACA 60  
 Db 199 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAAGCCAGATCAAGAACACA 258  
 QY 61 ACCCAACATACCTCACTCAGGATCTCAGCTTGGATCAGCTTCCCAATCTGTCGAA 120  
 Db 259 ACCCAACATACCTCACTCAGGATCTCAGCTTGGATCAGCTTCCCAATCTGTCGAA 318  
 QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACCAACCAAGGAGTCAAGTCAACCTG 180  
 Db 319 ATTACATCAACATCAACCACTACTAGCTTCAACCAACCAAGGAGTCAAGTCAACCTG 378  
 QY 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240  
 Db 379 CAATCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 438  
 QY 241 ACTACAAAACAGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 Db 439 ACCAACAACAGCGCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 498

QY 301 GTGTTAACTTTGTACCTTGCAGCATATGCGAGCAACAATCCAACTGCTGGCTATCTGC 360  
 Db 499 GTGTTCAACTTTGTACCTTGCAGCATATGCGAGCAACAATCCAACTGCTGGCTATCTGC 558  
 QY 361 AAAAGGAATACCAAAACAAAAACCCAGGAAGAAAAACCAACCAAGCCTACAAAAAACCA 420  
 Db 559 AAAAGGAATACCAAAACAAAAACCCAGGAAGAAAAACCACTACCAAGCCCAAAAAAACCA 618  
 QY 421 ACCTTCAAGACAAACCAAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCA 480  
 Db 619 ACCTTCAAGACAAACCAAAAAAGATCCCAAACTCAAACTCAAACTCAAACTCAAACTCA 678  
 QY 481 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 Db 679 ACCACCAAGCCCAAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
 QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 Db 739 CTACTCACTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
 QY 601 TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACCAACCAACCAACCAACCA 660  
 Db 799 TCAACTTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCTACCAACCAACCAACCAACCA 858  
 QY 661 TCACAACTCTCATCTCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 699  
 Db 859 TCACAACTCTCATCTCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 897  
 RESULT 14  
 AAV38298  
 ID AAV38298 standard; DNA; 696 BP.  
 XX AAV38298;  
 AC AAV38298;  
 DT 26-OCT-1998 (first entry)  
 DE Respiratory syncytial virus glycoprotein G gene portion.  
 KW RSV; glycoprotein G; pneumococcal surface protein A; PspA;  
 KW infection; Streptococcus pneumoniae; sepsis; otitis media;  
 KW meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;  
 KW ss.  
 XX Respiratory syncytial virus.  
 XX W09824927-AL.  
 XX 11-JUN-1998.  
 XX 04-DEC-1997; 97WO-US22847.  
 XX 04-DEC-1996; 96US-0759505.  
 XX (UYAL-) UNIV ALABAMA.  
 XX Briles DE, Curriel DT, McDaniel LS;  
 XX WPI; 1998-333343/29.  
 XX Plasmid containing pneumococcal epitope for expression in eukaryotic  
 XX cells - useful for eliciting immunological response to pneumococcal  
 XX infection or sepsis  
 XX Example 1; Fig 1C; 47pp; English.  
 XX This is a portion of the respiratory syncytial virus glycoprotein G  
 XX (RSVG) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)  
 XX to create plasmid pG74. This plasmid contains a human  
 XX cytomegalovirus immediate early promoter and the RSVG gene portion  
 XX such that when an in-frame fusion is made, the resultant fusion  
 XX protein may be transported to, and anchored in, a mammalian cell

CC membrane where it can be exposed to the host immune system.  
 CC Insertion of pneumococcal surface protein A (PspA) coding sequence  
 CC created plasmid pKSD2601. Intramuscular immunisation of BALB/c  
 CC mice with pKSD2601 induced protection against an otherwise lethal  
 CC challenge with a capsular type 3 pneumococcus. A claimed plasmid  
 CC for expression of pneumococcal epitope DNA in eukaryotic cells  
 CC includes a promoter for driving expression in a eukaryotic cell  
 CC (e.g., HCMV-IE), DNA encoding a leader sequence (e.g. of RSV)  
 CC and DNA encoding a pneumococcal epitope such as PspA. The invention  
 CC also provides a vaccine comprising the plasmid and a suitable  
 CC carrier or diluent, and optionally one or more cytokines or DNA  
 CC encoding them, or a bacterial delivery system. The vaccine is used  
 CC to elicit an immunological response in a host, including humans,  
 CC susceptible to pneumococcal infection or sepsis. The plasmid can  
 CC also be used to express a pneumococcal epitope of interest in vitro.  
 XX  
 SQ Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;  
 Query Match 62.3%; Score 445.8; DB 19; Length 696;  
 Best Local Similarity 95.3%; Pred. No. 8.4e-98;  
 Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;  
 QY 1 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60  
 Db 199 CACAAAGTCACACCAACCACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 258  
 QY 61 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGCTTCTCCAAATCTGTCTGAA 120  
 Db 259 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGCTTCTCCAAATCTGTCTGAA 318  
 QY 121 ATTACATCAAAACCAACCACTACTAGCTTCACACACACAGAGTCAAGTCAAACTTG 180  
 Db 319 ATTACATCAAAATCACCACCATCTAGCTTCACACACACAGAGTCAAGTCAAACTTG 378  
 QY 181 CAACCCACACAGTCACAGCTTAAACACACAAAC-AAACCCAAACACACACCAAGCC 239  
 Db 379 CAATCCACACAGTCACAGCTTAAACACACAAAC-AAACCCAAACACACCAAGCC 438  
 QY 240 CACTACAAACCAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 299  
 Db 439 CACCACAAACCAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
 QY 300 AGTGTCTTAACTTCTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCAACCA 359  
 Db 499 AGTGTCTTAACTTCTACCTGCAGATATGAGCAACCAACCAACCAACCAACCAACCAACCA 558  
 QY 360 CAAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 419  
 Db 559 CAAAGGAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
 QY 420 AACTTCAGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 479  
 Db 619 AACTTCAGACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
 QY 480 CACCACCAAGCCC 492  
 Db 679 CACCACCAAGCCC 691  
 RESULT 15  
 ID AAV17552 standard; cDNA; 15225 BP.  
 XX  
 AC AAV17552;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE Respiratory syncytial virus genome.  
 XX  
 KW RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.  
 XX  
 OS Human respiratory syncytial virus B-1.  
 XX

PN W09802530-A1.  
 XX  
 PD 22-JAN-1998.  
 XX  
 PF 15-JUL-1997; 97WO-US12269.  
 XX  
 PR 23-MAY-1997; 97US-0047634.  
 PR 15-JUL-1996; 96US-0021773.  
 PR 09-MAY-1997; 97US-0046141.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 PI Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;  
 PI Whitehead SS;  
 XX  
 DR WPI; 1998-110579/10.  
 XX  
 PT Attenuated respiratory syncytial virus vaccines - useful to protect  
 PT individuals against RSV infection  
 XX  
 PS Example 4; Page 195-202; 238pp; English.  
 XX  
 CC This is the complete nucleotide sequence of the wild-type B-1  
 CC respiratory syncytial virus (RSV). The genome is negative-sense;  
 CC the 5'-3' positive-sense sequence of D46 is provided in AAV17553.  
 CC A novel infectious recombinant RSV comprises a RSV genome or  
 CC antigenome, a major nucleocapsid (N) protein, a nucleocapsid  
 CC phosphoprotein (P), a large polymerase protein (L), and a RNA  
 CC polymerase elongation factor, where the recombinant RSV has at  
 CC least two attenuating mutations, one of the mutations specifying a  
 CC temperature-sensitive (ts) substitution at amino acid Phe521,  
 CC Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts  
 CC nucleotide substitution in the gene-start sequence of gene M2.  
 CC Also claimed are: (1) an isolated infectious RSV particle which  
 CC comprises a recombinant RSV (anti)genome, N, P, and L proteins, a  
 CC RNA polymerase elongation factor, where the (anti)genome is modified:  
 CC (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or  
 CC a cis-acting regulatory sequence; and (ii) by a termination codon  
 CC introduced within a selected gene; or by a change in sequence,  
 CC position or presence of a GS or GE transcription signal relative to  
 CC the selected gene; (2) an expression vector; and (3) an RSV strain  
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR  
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),  
 CC cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or  
 CC B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated  
 CC attenuated recombinant RSV and RSV particles are used in a vaccine  
 CC to stimulate the immune system of an individual to induce  
 CC protection against RSV. The expression vector of (2) is used for  
 CC the production of infectious attenuated RSV particles.  
 XX  
 SQ Sequence 15225 BP; 5904 A; 2710 C; 2399 G; 4212 T; 0 other;  
 Query Match 42.0%; Score 300.4; DB 19; Length 15225;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-62;  
 Matches 464; Conservative 0; Mismatches 251; Indels 3; Gaps 1;  
 QY 1 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60  
 Db 4888 CACAAAGTCACACTAACAACCTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 4947  
 QY 61 ACCCAACATACCTCACTCAGAGTCTCAGTTGGAATCAGCTTCTCCAAATCTGTCTGAA 120  
 Db 4948 ATCACCACCTTACTTACTCAAGTCCCACCAAGAGGGTTAGCTCATCCAAACCACTACA 5007  
 QY 121 ATTACATCAACACCAACCACTACTAGCTTCACACACACAGAGTCAAGTCAAACTTG 180  
 Db 5008 ACCATCATCCCAATCCACCAAAATTCAGCCACACATCATCCCAACCAAAAGTCAGAAACA 5067  
 QY 181 CAAACCCACACAGTCACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240  
 Db 5068 CACCAACACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5127  
 QY 241 ACTACAAAACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 16:51:32 ; Search time 2923.85 Seconds  
(without alignments)  
10004.074 Million cell updates/sec

Title: US-09-462-816-3  
Perfect score: 715  
Sequence: 1 cacaaagtcacactaacac.....gtagttattataaaaaaaa 715

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_man.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	715	100.0	715	6	BD081934
2	715	100.0	920	6	A16257
3	715	100.0	920	6	A16258
4	715	100.0	920	6	AR080406
5	715	100.0	920	6	AR092530
6	715	100.0	920	6	AR122885
7	715	100.0	920	6	AR123540
8	715	100.0	920	6	AR148357
9	715	100.0	920	6	BD081933
10	710	99.3	923	14	HRSRNAG
11	709	99.2	922	14	HRSVGL16
12	699.2	97.8	917	14	RSHGLYG
13	696	97.3	894	6	AR080424
14	696	97.3	894	6	AR092548
15	654.8	91.6	8510	14	RSHICE
16	654.8	91.6	15222	6	AR093219
17	654.8	91.6	15222	6	BD091237
18	654.8	91.6	15222	14	HRU50362
19	654.8	91.6	15222	14	HRU50363
20	654.8	91.6	15222	14	HRU63644
21	654.8	91.6	15222	14	RSHSEQ
22	654.8	91.6	15223	6	AR089137
23	654.8	91.6	15223	14	AF035006
24	649.6	90.9	914	14	PARSENVG
25	649.6	90.9	918	6	AX047075
26	646.2	90.4	897	6	AX339011
27	641.8	89.8	15190	14	RSU39662
28	641.8	89.8	15191	14	RSU39661
29	628.6	87.9	840	14	AF193308
30	628.2	87.9	918	14	AF065406
31	625.4	87.5	840	14	AF193306
32	621	86.9	922	14	HRSVGL12
33	621	86.9	922	14	HRSVGL11
34	621	86.9	922	14	HRSVGL10
35	621	86.9	922	14	HRSVGL09
36	620.8	86.8	887	14	AF065409
37	620.2	86.7	918	14	AF065405
38	619.4	86.6	922	14	HRSVGL2
39	618.4	86.5	922	14	HRSVGL1
40	618.4	86.5	922	14	HRSVGL0
41	618.4	86.5	922	14	HRSVGL0
42	617.8	86.4	922	14	AF065256
43	617.8	86.4	922	14	AF065258
44	617.8	86.4	922	14	HRSVGL13
45	617.8	86.4	922	14	HRSVGL9

ALIGNMENTS

RESULT 1	BD081934	715 bp	DNA	linear	PAT 27-AUG-2002
BD081934	Nucleic acid vaccines encoding G protein of respiratory syncytial virus.				
LOCUS	BD081934.1	GI:22627544			
DEFINITION	Respiratory syncytial virus				
ACCESSION	BD081934				
VERSION	JP 2001512662-A/2.				
KEYWORDS	Respiratory syncytial virus				
SOURCE	Respiratory syncytial virus				
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.				
REFERENCE	1 (bases 1 to 715)				
AUTHORS	Li, X.; Sambhara, S. and Klein, M.H.				
TITLE	Nucleic acid vaccines encoding G protein of respiratory syncytial				

JOURNAL	Patent: JP 2001512662-A 2 28-AUG-2001;
COMMENT	CONNAUGHT LABORATORIES LTD OS Respiratory syncytial virus PN JP 2001512662-A/2 PD 28-AUG-2001 PR 16-JUL-1998 JP 2000503216 PR 18-JUL-1997 US 08/896442 PI XIAOMAO LI,SURYAPRAKASH SAMBHARA,MICHEL H KLEIN PC C12N15/09,A61K48/00,C07K16/10,C12P21/08,C12Q1/68,G01N33/53, PC C12N15/00 CC Nucleic acid vaccines encoding G protein of respiratory CC syncytial virus
FEATURES	FH Key Location/Qualifiers FT CDS (1)..(699).
source	1..715 /organism="Respiratory syncytial virus" /mol_type="genomic DNA" /db_xref="taxon:12814"
BASE COUNT	305 a 240 c 70 g 100 t
ORIGIN	
Query Match	100.0%; Score 715; DB 6; Length 715;
Best Local Similarity	100.0%; Pred. No. 3.3e-151;
Matches 715; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1	CACAAAGTCACACTAAACAACATGCAATCATATACAAAGATGCAACAAGCCAGATCAAGAAACACA 60
DB 1	CACAAAGTCACACTAAACAACATGCAATCATATACAAAGATGCAACAAGCCAGATCAAGAAACACA 60
QY 61	ACCCACATACCTCAGCTCAGGATCCTCAGCTTGGATCAGCTTCTCCAAATCTGTCGAA 120
DB 61	ACCCACATACCTCAGCTCAGGATCCTCAGCTTGGATCAGCTTCTCCAAATCTGTCGAA 120
QY 121	ATTATCATCAAAAACCCACCATACTAGCTTCAACAACACAGAGAGTCAAGTCAAAACCTG 180
DB 121	ATTATCATCAAAAACCCACCATACTAGCTTCAACAACACAGAGAGTCAAGTCAAAACCTG 180
QY 181	CAACCCCAACAGCTCAAGACTTAAAAACAACAACAAACCCAAAAACAACCCAGCAAGGCC 240
DB 181	CAACCCCAACAGCTCAAGACTTAAAAACAACAACAAACCCAAAAACAACCCAGCAAGGCC 240
QY 241	ACTACAAAAACCGCCAAAAACAACCCAAAAACAACCCAAATATGATTTTCATCTTGAA 300
DB 241	ACTACAAAAACCGCCAAAAACAACCCAAAAACAACCCAAATATGATTTTCATCTTGAA 300
QY 301	GTGTTTAACTTTGTACCTCGAGCATATGCAGCAACCAATCCAACTGCTGGGCTATCTGC 360
DB 301	GTGTTTAACTTTGTACCTCGAGCATATGCAGCAACCAATCCAACTGCTGGGCTATCTGC 360
QY 361	AAAAGATATCCAAAAAACAACCCAGGAAGAAAACCCACCAACAGCCCTTACAAAAAACCA 420
DB 361	AAAAGATATCCAAAAAACAACCCAGGAAGAAAACCCACCAACAGCCCTTACAAAAAACCA 420
QY 421	ACCTTCAAGCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCCAAAGGAAGTACCC 480
DB 421	ACCTTCAAGCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCCAAAGGAAGTACCC 480
QY 481	ACCACCAAGCCCAAGAGAGCCCAACCATCAACCAACCCAAAAACAACATCACAACTACA 540
DB 481	ACCACCAAGCCCAAGAGAGCCCAACCATCAACCAACCCAAAAACAACATCACAACTACA 540
QY 541	CTGCTCACCAACAACACACACAGGAAATCCAAAACTCAACAGTCAAAATGGAACCTTCCAC 600
DB 541	CTGCTCACCAACAACACACAGGAAATCCAAAACTCAACAGTCAAAATGGAACCTTCCAC 600
QY 601	TCAACCTCTCCGAAGGCAATCTAAGGCCCTTCTCAAGTCTCCAAATCCGAGCAACCCA 660
DB 601	TCAACCTCTCCGAAGGCAATCTAAGGCCCTTCTCAAGTCTCCAAATCCGAGCAACCCA 660
QY 661	TCACAAAGCCTCATCTCCACCCCAACAACAGCCAGTATGTTATTAATAAAAAA 715
DB 661	TCACAAAGCCTCATCTCCACCCCAACAACAGCCAGTATGTTATTAATAAAAAA 715

[illegible]











/db\_xref="SPTREMBL:Q01929"  
/translation="MSKNKQDQRTAKTLEKTKWTDLNLLFISSGLYKLNLSIAQITLS  
ILAMIISTSLIITAIIFASANKHVTITAIQDATSOIKNTTPTLTQDPQLGISFS  
NLSEITSTVITLSTTPGVSKNLQPTTKNTTKNTTQPSKPTTKQRQKNPKPN  
NDPHEVFNFPVPCISNNPFCWAICRIIPNKKPGKTTTATPKPTFKTKDLKPKQ  
TKKEVPTTKPTTEPTTNTTKNTTITLLTNTTGNPKLTSQMETFHTSSBGNLSP  
SQVSTTSEHPSPSPSPNTTQ"

BASE COUNT	379 a	290 c	98 g	156 t
ORIGIN				
Query Match	99.3%;	Score 710;	DB 14;	Length 923;
Best Local Similarity	100.0%;	Pred. No. 4.4e-150;		
Matches 710;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACCAAGCCAGATCAAGAACACA	60	
Db	214	CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACCAAGCCAGATCAAGAACACA	273	
QY	61	ACCCCAACATACCTCACTCAGAGTCTCTCAGTTTGGAAATCAGCTTCTCCAAATCTGTCTGAA	120	
Db	274	ACCCCAACATACCTCACTCAGAGTCTCTCAGTTTGGAAATCAGCTTCTCCAAATCTGTCTGAA	333	
QY	121	ATTACATCACAACCAACCACTACTAGCTTCAACACACCAAGAGTCAAGTCAAACTG	180	
Db	334	ATTACATCACAACCAACCACTACTAGCTTCAACACACCAAGAGTCAAGTCAAACTG	393	
QY	181	CAACCCCAACAGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	240	
Db	394	CAACCCCAACAGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	453	
QY	241	ACTACAAAACAAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA	300	
Db	454	ACTACAAAACAAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA	513	
QY	301	GTGTTTAACTTTGTACCTTGCAGCATATGTCAGCAACCAATCCAACTCTGGGCTATCTGC	360	
Db	514	GTGTTTAACTTTGTACCTTGCAGCATATGTCAGCAACCAATCCAACTCTGGGCTATCTGC	573	
QY	361	AAAAGAAATACCAACAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA	420	
Db	574	AAAAGAAATACCAACAAACCAAGGAAAGAAACCAACCAACCAACCAACCAACCAACCA	633	
QY	421	ACCTTCAAGCAACCAACCAACCAAGGAAATCTCAAACTCAAACTCAAACTCAAACTCA	480	
Db	634	ACCTTCAAGCAACCAACCAACCAAGGAAATCTCAAACTCAAACTCAAACTCAAACTCA	693	
QY	481	ACCACCAAGCCCAAGAGGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCA	540	
Db	694	ACCACCAAGCCCAAGAGGAGCCCAACCATCAACCAACCAACCAACCAACCAACCAACCA	753	
QY	541	CTGCTCAACCAACCAACCAAGGAAATCTCAAACTCAAACTCAAACTCAAACTCAAACT	600	
Db	754	CTGCTCAACCAACCAACCAAGGAAATCTCAAACTCAAACTCAAACTCAAACTCAAACT	813	
QY	601	TCAACCTCTCCAGGCAATTAAGCCCTTCTCAAGTCTCCAACTCCAACTCCAGGAGCCCA	660	
Db	814	TCAACCTCTCCAGGCAATTAAGCCCTTCTCAAGTCTCCAACTCCAACTCCAGGAGCCCA	873	
QY	661	TCACACCTCTCACTCCACCAACCAACCAAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGT	710	
Db	874	TCACACCTCTCACTCCACCAACCAACCAAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGT	923	

RESULT 11  
HRSVGL16  
LOCUS  
DEFINITION  
Human respiratory syncytial virus (MON-7-91) subgroup A, G gene for  
Glycoprotein.  
ACCESSION  
VERSION  
233429.1  
GI:485888  
KEYWORDS  
G gene; glycoprotein.  
SOURCE  
Human respiratory syncytial virus  
ORGANISM  
Human respiratory syncytial virus

Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Pneumovirinae; Pneumovirus.  
1 (bases 1 to 922)  
Garcia, O., Martin, M., Dopazo, J., Arbiza, J., Fabraile, S., Russi, J.,  
Hortal, M., Perez-Brena, P., Martinez, I., Garcia-Barreno, B. and  
Melero, J.A.  
Evolutionary pattern of human respiratory syncytial virus (subgroup  
A): cocirculating lineages and correlation of genetic and antigenic  
changes in the G glycoprotein  
J. Virol. 68 (9), 5448-5459 (1994)  
94335057  
8057427  
2 (bases 1 to 922)  
Dopazo, J.  
Direct Submission  
Submitted (10-MAY-1994) Dopazo J., Centro Nacional de Biotecnologia  
- CSIC, Bioinformatica, Universidad Autonoma, Cantoblanco, Madrid,  
SPAIN, 28049  
Location/Qualifiers  
1..922  
/organism="Human respiratory syncytial virus"  
/mol\_type="genomic RNA"  
/strain="subgroup A"  
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"  
/db\_xref="taxon:11250"  
/cell\_line="HEp-2"  
16..912  
/gene="G"  
16..912  
/partial  
/gene="G"  
/codon\_start=1  
/product="glycoprotein"  
/protein\_id="CAA83872.1"  
/db\_xref="GI:485889"  
/db\_xref="SPTREMBL:Q82068"  
/translation="ISKNDQRTAKTLEKTKWTDLNLLFISSGLYKLNLSIAQITLS  
NLSEITSTVITLSTTPGVSKNLQPTTKNTTKNTTQPSKPTTKQRQKNPKPN  
NDPHEVFNFPVPCISNNPFCWAICRIIPNKKPGKTTTATPKPTFKTKDLKPKQ  
TKKEVPTTKPTTEPTTNTTKNTTITLLTNTTGNPKLTSQMETFHTSSBGNLSP  
SQVSTTSEHPSPSPSPNTTQ"

BASE COUNT 378 a 291 c 97 g 156 t  
ORIGIN

Query Match 99.2%; Score 709; DB 14; Length 922;  
Best Local Similarity 100.0%; Pred. No. 7.4e-150;  
Matches 709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACCAAGCCAGATCAAGAACACA 60  
Db 214 CACAAAGTCACACTAAACAACTGCAATCATACAGATGCAACCAAGCCAGATCAAGAACACA 273  
QY 61 ACCCAACATACCTCACTCAGGATCTCTCAGTTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 120  
Db 274 ACCCAACATACCTCACTCAGGATCTCTCAGTTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 333  
QY 121 ATTACATCACAACCAACCACTACTAGCTTCAACACCAAGAGTCAAGTCAAACTG 180  
Db 334 ATTACATCACAACCAACCACTACTAGCTTCAACACCAAGAGTCAAGTCAAACTG 393  
QY 181 CAACCCCAACAGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240  
Db 394 CAACCCCAACAGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 453  
QY 241 ACTACAAAACAAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
Db 454 ACTACAAAACAAAGCCCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513  
QY 301 GTGTTTAACTTTGTACCTTGCAGCATATGTCAGCAACCAATCCAACTCTGGGCTATCTGC 360  
Db 514 GTGTTTAACTTTGTACCTTGCAGCATATGTCAGCAACCAATCCAACTCTGGGCTATCTGC 573

361 AAAAGAATACCAACCAAAAAACCGAAGAAAGAAACCAACCAAGCCCTTACAAAAAACCA 420  
 574 AAAAGAATACCAACCAAAAAACCGAAGAAAGAAACCAACCAAGCCCTTACAAAAAACCA 633  
 421 ACCTTCAAGACACCAAAAAAGATCTCAACCTCAACCACTAAACCAAGGAAGTACCC 480  
 634 ACCTTCAAGACACCAAAAAAGATCTCAACCTCAACCACTAAACCAAGGAAGTACCC 693  
 481 ACCACCAAGCCACAGAAGAGCCACCAATCAACCAACCAACCAACCAACCAACTACA 540  
 694 ACCACCAAGCCACAGAAGAGCCACCAATCAACCAACCAACCAACCAACCAACTACA 753  
 541 CTGCTCAACCAACCAACCAAGGAATCCAAATCTCAACAGTCAATGGAACCTTCCAC 600  
 754 CTGCTCAACCAACCAACCAAGGAATCCAAATCTCAACAGTCAATGGAACCTTCCAC 813  
 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 660  
 814 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 873  
 661 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 709  
 874 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 922

RESULT 12  
 RSHGLYG 917 bp ss-RNA linear VRL 03-AUG-1993  
 LOCUS Human respiratory syncytial virus (subgroup A) attachment protein  
 DEFINITION (G) mRNA, complete cds.  
 MI7212  
 VERSION 1 GI:333940  
 KEYWORDS attachment glycoprotein; surface glycoprotein.  
 SOURCE Human respiratory syncytial virus  
 ORGANISM Human respiratory syncytial virus  
 Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 1 (bases 1 to 917)  
 Johnson, P.R., Spriggs, M.K., Olmsted, R.A., and Collins, P.L.  
 The G glycoprotein of human respiratory syncytial viruses of  
 subgroups A and B: extensive sequence divergence between  
 antigenically related proteins  
 Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)  
 87289657  
 2441388  
 Original source text: Human respiratory syncytial virus (subgroup A,  
 strain long), cDNA to viral RNA, clones pUF63, pUB69, and pA3.  
 The exact 5' end of long G mRNA was not determined.

FEATURES  
 source  
 1. .917  
 /organism="Human respiratory syncytial virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:11250"  
 16. .912  
 /note="Attachment glycoprotein (G)"  
 /codon\_start=1  
 /protein\_id="AAA47411.1"  
 /db\_xref="GI:333941"  
 /translation="MSKNDQRTAKLTLEKWTDLNHLFTSSGLYKLNKLSIAQITLS  
 ILAMISTLIITAIPIASANKVLTITAIIDATSIQKNTTPTLYLTOPQLGISFS  
 NLSIITSQTTIILASTPPGVKSLQPTTKVTKNTTQTQPSKPTTKQRONKPKPN  
 NDFHFEVFNPCISGNSNPTCWAICRIPNKPGKTKTKPKTKPTTKTKDHPQ  
 TTKPKVPTTKPTTEPTINTKTNIIITLLTNTTGNPKLTQSMETPHSTSSGNLSP  
 SQVSTSEHSPQSPPNTTQ"  
 BASE COUNT 375 a 290 c 97 g 154 t 1 others  
 ORIGIN  
 Query Match 97.8%; Score 699.2; DB 14; Length 917;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-147;  
 Matches 701; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 1 CACAAAGTCACTAACCACTGCAATCATACAGATCAACAGCCAGATCAAGAACCA 60  
 |||||

214 CACAAAGTCACTAACCACTGCAATCATACAGATCAACAGCCAGATCAAGAACCA 273  
 61 ACCCAACATACCTCTACTAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 120  
 274 ACCCAACATACCTCTACTAGGATCTCTAGCTTGGAAATCAGCTTCTCCAAATCTGTCTGAA 333  
 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACACACAGAGTCAAGTCAAGCTG 180  
 334 ATTACATCAACAAACCAACCACTACTAGCTTCAACACACAGAGTCAAGTCAAGCTG 393  
 181 CAACCCCAACAGTCAAGACTTAAAAACACAAACCAACCCCAACCAACCAACCAAGCC 240  
 394 CAACCCCAACAGTCAAGACTTAAAAACACAAACCAACCCCAACCAACCAACCAAGCC 453  
 241 ACTACAAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300  
 454 ACTACAAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513  
 301 GTGTTTAACTTTGTACCTCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 360  
 514 GTGTTTAACTTTGTACCTCTGAGCATATGAGCAACCAACCAACCAACCAACCAACCA 573  
 361 AAAAGATATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 574 AAAAGATATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633  
 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 634 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693  
 481 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 694 ACCACCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753  
 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 754 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 813  
 601 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 660  
 814 TCAACCTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACCAATCCGAGACCCA 873  
 661 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 704  
 874 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 917

RESULT 13  
 AR080424  
 LOCUS 894 bp DNA linear PAT 31-AUG-2000  
 DEFINITION Sequence 28 from patent US 5968776.  
 AR080424  
 ACCESSION AR080424  
 VERSION AR080424.1 GI:10007159  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 1 (bases 1 to 894)  
 Klein, M.H., Du, R.-P. and Ewasysghyn, M.R.  
 Multimeric hybrid gene encoding a chimeric protein which confers  
 protection against parainfluenza virus and respiratory syncytial  
 virus  
 Patent: US 5968776-A 28 19-OCT-1999;  
 JOURNAL Location/Qualifiers  
 FEATURES  
 1. .894  
 /organism="unknown"  
 BASE COUNT 364 a 288 c 93 g 149 t  
 ORIGIN  
 Query Match 97.3%; Score 696; DB 6; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-147;  
 Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAACCACTGCAATCATACAGAGTGCACCAAGCCAGATCAAGAACACA 60  
 DB 199 CACAAAGTCACACTAACCACTGCAATCATACAGAGTGCACCAAGCCAGATCAAGAACACA 258  
 QY 61 ACCCCAAACATACCTCAGTCAGATCCTCAGCTTGGATCAGCTTCTCCATCTGCTCGAA 120  
 DB 259 ACCCCAAACATACCTCAGTCAGATCCTCAGCTTGGATCAGCTTCTCCATCTGCTCGAA 318  
 QY 121 ATTACATCACAAACACCAACCACTATAGCTTCAACCAACAGGAGTCAAGTCAAACTG 180  
 DB 319 ATTACATCACAAACACCAACCACTATAGCTTCAACCAACAGGAGTCAAGTCAAACTG 378  
 QY 181 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
 DB 379 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438  
 QY 241 ACTACAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 300  
 DB 439 ACTACAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 498  
 QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTGCTGGGCTATCTGC 360  
 DB 499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTGCTGGGCTATCTGC 558  
 QY 361 AAGAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 DB 559 AAGAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 DB 619 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
 QY 481 ACACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB 679 ACACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
 QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 DB 739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
 QY 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCCA 660  
 DB 799 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCCA 858  
 QY 661 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCAG 696  
 DB 859 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCAG 894

RESULT 14  
 AR092548  
 LOCUS  
 DEFINITION Sequence 28 from patent US 5998169.  
 ACCESSION AR092548  
 VERSION AR092548.1 GI:10019302  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS Klein,M.H., Du,R.-P. and Ewaszyshyn,M.E.  
 TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus  
 JOURNAL Patent: US 5998169-A 28 07-DEC-1999;  
 FEATURES Location/Qualifiers  
 source 1..894  
 /organism="unknown"  
 BASE COUNT 364 a 288 c 93 g 149 t  
 ORIGIN  
 Query Match 97.3%; Score 696; DB 6; Length 894;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-147;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CACAAAGTCACACTAACCACTGCAATCATACAGAGTGCACCAAGCCAGATCAAGAACACA 60  
 DB 199 CACAAAGTCACACTAACCACTGCAATCATACAGAGTGCACCAAGCCAGATCAAGAACACA 258  
 QY 61 ACCCCAAACATACCTCAGTCAGATCCTCAGCTTGGATCAGCTTCTCCATCTGCTCGAA 120  
 DB 259 ACCCCAAACATACCTCAGTCAGATCCTCAGCTTGGATCAGCTTCTCCATCTGCTCGAA 318  
 QY 121 ATTACATCACAAACACCAACCACTATAGCTTCAACCAACAGGAGTCAAGTCAAACTG 180  
 DB 319 ATTACATCACAAACACCAACCACTATAGCTTCAACCAACAGGAGTCAAGTCAAACTG 378  
 QY 181 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
 DB 379 CAAACCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438  
 QY 241 ACTACAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 300  
 DB 439 ACTACAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 498  
 QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTGCTGGGCTATCTGC 360  
 DB 499 GTGTTTAACTTTGTACCTGAGCATATGAGCAACCAATCCAACTGCTGGGCTATCTGC 558  
 QY 361 AAGAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420  
 DB 559 AAGAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618  
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480  
 DB 619 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678  
 QY 481 ACACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 540  
 DB 679 ACACCAAGCCACAGAGAGCCACCATCAACCAACCAACCAACCAACCAACCAACCAACCA 738  
 QY 541 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600  
 DB 739 CTGCTCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798  
 QY 601 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCCA 660  
 DB 799 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCCA 858  
 QY 661 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCAG 696  
 DB 859 TCACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCAACTCCGAGCAGCCAG 894

RESULT 15  
 RSHICE  
 LOCUS  
 DEFINITION Human respiratory syncytial virus nonstructural protein (1C), nonstructural protein (1B), major nucleocapsid (N), phosphoprotein (P), protein (M), 1A (1A), G (G), protein (P) and envelope-associated protein (22K) gene, complete cds.  
 ACCESSION M11486 K01459 K02719 K03348 K03349 M11217 M11244 M11487 M11505  
 VERSION M11514 M11631 M12966  
 KEYWORDS M11486.1 GI:333925  
 SOURCE Human respiratory syncytial virus  
 ORGANISM Human respiratory syncytial virus  
 REFERENCE 1 (bases 1085 to 2220; 5254 to 5544)  
 AUTHORS Parameyoviridae; Pneumovirinae; Pneumovirus.  
 TITLE Elango,N. and Venkatesan,S.  
 JOURNAL Amino acid sequence of human respiratory syncytial virus nucleocapsid protein  
 Nucleic Acids Res. 11 (17), 5941-5951 (1983)



```
/gene="1B"
/replace="a"
1081..2277
/gene="N"
1081..2277
/gene="N"
1096..2271
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/protein_id="AAB59852.1"
/db_xref="GI:333928"
/translation="MALSKVINDLTKNQDLSKVTIORSTGDSIDTENVYQKHI
NKLGMILLITEDANHKFTGLIGLYAMSRIGREDTIKILRDAGVHYKANGVDVTHHQ
DINGKMKFVLTALSTETIQLINIEJESKSYKMKLEMEVAPETRHSDPCGMII
LCIAALVITKLAAGDRGLTAVIRRANNVLNEMKRYKGLLPDKDIANSFYEVFEKPHH
FIDVFVHFGIAQSVTRSGSRVEGIFAGLFMAYGAGQVLMFLWGLAKSVKNIMLGHAS
VQAMEQVEYEAQKLGGEAGFYHILNPKASLLSITOPPHFSSVVLGNAAGLGM
GEYRGTPRNQDLYDAAKAYARQLKENGVINYSVLDTAELEIAIKHQLNPKNDVEL"
2289..3196
/gene="P"
2289..3196
/gene="P"
2302..3027
/gene="P"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB59853.1"
/db_xref="GI:333929"
/translation="MEKFAPEFHGEDANNRATKFLIESIKGFTSPDKPKKDSIISVN
SIDIEVTKESPITSNTIINPTNEDDTAGNKPNYQRKPLVSKFEDTPSDNPFSLY
KETIETPDNNEESSYVEIENDQTNIDNITARDRIDELKSELGLMLHTLVASAGPT
SARDGIRDAMIGUREEMIEKIRTEALMTNDRLEMRNESEKMAKDTSDVSLNP
```

```
Query Match 91.6%; Score 654.8; DB 14; Length 8510;
Best local Similarity 94.8%; Pred. No. 1.1e-137;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAGTCACATACTGCAATCATATCAAGATGCAACAGCCAGATCAAGAACACA 60
Db 4839 CACAAGTCACACCAACTGCAATCATATCAAGATGCAACAGCCAGATCAAGAACACA 4898

QY 61 ACCCCACATACCTCACTCAGGATCCTCAGCTTGGATCAGTCTCCCAATCTGTCTGAA 120
Db 4899 ACCCCACATACCTCACTCAGGATCCTCAGCTTGGATCAGTCTCCCAATCTGTCTGAA 4958

QY 121 ATTACATCAAAACCAACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 180
Db 4959 ATTACATCAAAATCAACACCACTACTAGCTTCAACACACAGGAGTCAAGTCAAAACCTG 5018

QY 181 CAACCCACACAGTCACAGCTTAAACACACACACACACACACACACACACACACAC 240
Db 5019 CAATCCACACAGTCACAGCTTAAACACACACACACACACACACACACACACACAC 5078

QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
Db 5079 ACCACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5138

QY 301 GTGTTTAACTTTGTACCTCGAGCATATGCGACAACTCAACCTGCTGGGCTATCTGC 360
Db 5139 GTGTTTAACTTTGTACCTCGAGCATATGCGACAACTCAACCTGCTGGGCTATCTGC 5198

QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
Db 5199 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5258

QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
Db 5259 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5318

QY 481 ACCACCAAGCCACAGAGGACCAACCATCAACACCAACCAACCAACCAACCAACCA 540
```

```
Db 5319 ACCACCAAGCCACAGAGGACCAACCATCAACACCAACCAACCAACCAACCAACCA 5378
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
Db 5379 CTACTCACTCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5438
QY 601 TCACCTCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCCAACCAACCAACCA 660
Db 5439 TCAACTTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCTCAACATCCGAGTACCA 5498
QY 661 TCACACCTCTCATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCA 714
Db 5499 TCACACCTCTCATCTCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCA 5552
```

Search completed: October 30, 2003, 00:01:08  
Job time : 2924.85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:32:25 ; Search time 54.717 Seconds  
(without alignments)  
1094.143 Million cell updates/sec

Title: US-09-462-816-4  
Perfect score: 1229  
Sequence: 1 HKVLTALIQDARSQIKNT.....VSTTSEHPQSPPPTTRQ 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1229	100.0	298	12 Q82068	Q82068 human respi
2	1229	100.0	298	12 Q01929	Q01929 human respi
3	1033	88.9	298	12 Q09719	Q09719 respiratory
4	1093	88.9	298	12 Q09634	Q09634 human respi
5	1088	88.5	279	12 Q09678	Q09678 human respi
6	1072	87.2	279	12 Q096U0	Q096U0 human respi
7	1067	86.8	297	12 Q9YVB5	Q9YVB5 human respi
8	1066	86.7	297	12 Q82066	Q82066 human respi
9	1063	86.5	293	12 Q9YVB3	Q9YVB3 human respi
10	1062	86.4	298	12 Q9YVB4	Q9YVB4 human respi
11	1056	85.9	297	12 Q82071	Q82071 human respi
12	1055	85.8	278	12 Q09673	Q09673 human respi
13	1053	85.7	292	12 Q9YVB2	Q9YVB2 human respi
14	1053	85.7	297	12 Q82058	Q82058 human respi
15	1052	85.6	295	12 Q86359	Q86359 respiratory
16	1047	85.2	297	12 Q82067	Q82067 human respi

17	1047	85.2	297	12	Q81947	Q81947 human respi
18	1045	85.0	278	12	Q906T1	Q906T1 human respi
19	1045	85.0	295	12	Q86360	Q86360 respiratory
20	1044	84.9	279	12	Q9Q680	Q9Q680 human respi
21	1044	84.9	279	12	Q9Q6R9	Q9Q6R9 human respi
22	1044	84.9	293	12	Q9YVB1	Q9YVB1 human respi
23	1043	84.9	298	12	Q82065	Q82065 human respi
24	1041	84.7	278	12	Q9Q6U1	Q9Q6U1 human respi
25	1040	84.6	297	12	Q82074	Q82074 human respi
26	1040	84.6	297	12	Q82064	Q82064 human respi
27	1039	84.5	279	12	Q9Q6S2	Q9Q6S2 human respi
28	1039	84.5	292	12	Q9YVB0	Q9YVB0 human respi
29	1037	84.4	279	12	Q9Q6S9	Q9Q6S9 human respi
30	1037	84.4	295	12	Q86361	Q86361 respiratory
31	1037	84.4	297	12	Q82063	Q82063 human respi
32	1036	84.3	278	12	Q9Q6T4	Q9Q6T4 human respi
33	1036	84.3	279	12	Q9Q6S6	Q9Q6S6 human respi
34	1035	84.2	279	12	Q9Q6T9	Q9Q6T9 human respi
35	1035	84.2	299	12	Q82077	Q82077 human respi
36	1033	84.1	279	12	Q9Q6S7	Q9Q6S7 human respi
37	1033	84.1	295	12	Q86357	Q86357 respiratory
38	1032	84.0	279	12	Q9Q6S1	Q9Q6S1 human respi
39	1032	84.0	297	12	Q9YVC8	Q9YVC8 human respi
40	1032	84.0	297	12	Q91946	Q91946 human respi
41	1031	83.9	298	12	Q82060	Q82060 human respi
42	1029	83.7	297	12	Q82072	Q82072 human respi
43	1028	83.6	297	12	Q82079	Q82079 human respi
44	1028	83.6	297	12	Q82057	Q82057 human respi
45	1028	83.6	298	12	Q82062	Q82062 human respi

#### ALIGNMENTS

#### RESULT 1

Q82068 ID Q82068 PRELIMINARY; PRT; 298 AA.  
AC Q82068;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Glycoprotein (Fragment).  
GN G.  
OS Human respiratory syncytial virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11250;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Subgroup A;  
RX MEDLINE=94335057; PubMed=8057427;  
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisale S., Russi J.,  
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
RA Malero J.A.;  
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
RT A): cocirculating lineages and correlation of genetic and antigenic  
RT changes in the G glycoprotein.";  
RL J. Virol. 68:5448-5459(1994).  
DR EMBL; Z33429; CAA83872.1; -  
DR InterPro; IPR000925; Glycoprot G.  
DR InterPro; IPR006162; Ppantne attach.  
DR Pfam; PF00802; Glycoprotein G; 1.  
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.  
FT NON TER 1 1  
SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;  
Best local Similarity 100.0%; Pred. No. 3.8e-89;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTALIQDARSQIKNTPTLYTQDPQLGFSFNSLSEITSTQTTILASTPGVKSNI 60  
|||||  
DB 67 HKVLTALIQDARSQIKNTPTLYTQDPQLGFSFNSLSEITSTQTTILASTPGVKSNI 126



```

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OY 61 OPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 180
DB 187 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 2
Q01929 PRELIMINARY; PRT; 298 AA.
AC Q01929;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE G protein.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91065351; PubMed=2249671;
RA Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;
RT "Frame shift mutations as a novel mechanism for the generation of
RT newralization resistant mutants of human respiratory syncytial
RT virus.";
RL EMBO J. 9:4181-4187(1990).
DR EMBL; X17085; CAA34937.1; -
DR InterPro; IPR000925; Glycoprot G.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
SQ SEQUENCE 298 AA; 32771 MW; F8CD4213D97C2952 CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;
Best Local Similarity 100.0%; Pred. No. 3.8e-89; Mismatches 0; Indels 0; Gaps 0;
Matches 232; Conservative 0;

OY 1 HKVLTFTAIQDQTSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTTILASTTPGVKSNL 60
DB 67 HKVLTFTAIQDQTSQIKNTTPTLTQDPQLGIFSLSLSEITSTQTTTILASTTPGVKSNL 126
OY 61 OPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 180
DB 187 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 3
O09719 PRELIMINARY; PRT; 298 AA.
AC O09719;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Attachment glycoprotein (G).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

```

```

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OY 61 OPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120
DB 127 QPTVTKNTTTTQTPSKPTTKORQKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186
OY 121 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 180
DB 187 KRIPNKKPGKTTTKPTTKPTTKKPTTKKOLKPTTKPKKEVPTTKPTKEPTINTTKINITT 246
OY 181 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSSEGNLSPSOVSTTSEHPSQSPSPNTTRQ 298

RESULT 4
O09634 PRELIMINARY; PRT; 298 AA.
AC O09634;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Attachment protein (G).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Flows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S2;
RX MEDLINE=97185152; PubMed=9032893;
RA Tolley K.P., Marriott A.C., Simpson A., Flows D.J., Matthews D.A.,
RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
RA Pringle C.R.;
RT "Identification of mutations contributing to the reduced virulence of
RT a modified strain of respiratory syncytial virus.";
RL Vaccine 14:1637-1646(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S2;
RA Easton A.J.;
RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

```

```
DR EMBL; U39662; AAC57026.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppanthe attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match      88.9%; Score 1093; DB 12; Length 298;
Best Local Similarity 90.1%; Pred. No. 1.9e-78;
Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPGVKSNL 60
DB 67 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPGVKSNL 126

QY 61 QPTTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 127 QSTTVTKNTNTTQIQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186

QY 121 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 180
DB 187 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 246

QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 232
DB 247 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 298

RESULT 5
Q906T8 PRELIMINARY; PRT; 279 AA.
AC Q906T8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=92011;
RX MEDLINE=20283719; PubMed=10823752;
RA Choi E.H., Lee H.J.;
RT "Genetic diversity and molecular epidemiology of the G protein of
RT subgroups A and B of respiratory syncytial virus isolated over 9
RT consecutive epidemics in Korea.";
RL J. Infect. Dis. 181:1547-1556(2000).
DR EMBL; AF193308; AAF23731.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppanthe attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
FT NON TER
SQ SEQUENCE 279 AA; 30498 MW; 4E1CDF2F7E569E3A CRC64;

Query Match      88.5%; Score 1088; DB 12; Length 279;
Best Local Similarity 90.1%; Pred. No. 4.4e-78;
Matches 209; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPGVKSNL 60
DB 48 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPSVSTL 107

QY 61 QPTTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 108 LPTTVTKNTNTTQIQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 167

QY 121 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 180
DB 168 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 227

QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 230
DB 228 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 277

RESULT 7
Q91VB5 PRELIMINARY; PRT; 297 AA.
AC Q91VB5;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99022964; PubMed=9806017;
RX MEDLINE=99022964; PubMed=9806017;
```

```
QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 232
DB 228 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTTQ 279

RESULT 6
Q906U0 PRELIMINARY; PRT; 279 AA.
AC Q906U0;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91242;
RX MEDLINE=20283719; PubMed=10823752;
RA Choi E.H., Lee H.J.;
RT "Genetic diversity and molecular epidemiology of the G protein of
RT subgroups A and B of respiratory syncytial virus isolated over 9
RT consecutive epidemics in Korea.";
RL J. Infect. Dis. 181:1547-1556(2000).
DR EMBL; AF193306; AAF23729.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppanthe attach.
DR Pfam; PF00802; Glycoprotein G; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
FT NON TER
SQ SEQUENCE 279 AA; 30544 MW; 6B270AAEB1CB3533 CRC64;

Query Match      87.2%; Score 1072; DB 12; Length 279;
Best Local Similarity 90.0%; Pred. No. 8e-77; Indels 0; Gaps 0;
Matches 207; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPGVKSNL 60
DB 48 HKVLTLLTALIQDQATSOIKNTTPTVLTQDPQLGIFSLSNLSSETTSQTTLILASTTPSVSTL 107

QY 61 QPTTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120
DB 108 LPTTVTKNTNTTQIQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 167

QY 121 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 180
DB 168 KRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 227

QY 181 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 230
DB 228 LTNNTTGNPKLTQSMETFHSTSSGNSLSPSQVSTTSEHPSQSPSPNTT 277

RESULT 7
Q91VB5 PRELIMINARY; PRT; 297 AA.
AC Q91VB5;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Attachment glycoprotein G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99022964; PubMed=9806017;
RX MEDLINE=99022964; PubMed=9806017;
```

RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;  
 RT "Antigenic and genetic diversity among the attachment proteins of  
 RT group A respiratory syncytial viruses that have caused repeat  
 RT infections in children.";  
 RL J. Infect. Dis. 178:925-932 (1998).  
 DR EMBL; AF065405; AAD02941.1; -;  
 DR InterPro; IPR000925; Glycoprot G.  
 DR InterPro; IPR006162; Ppantne attach.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
 SQ SEQUENCE 237 AA; 32572 MW; 5C10E719A05EF5C1 CRC64;

Query Match 86.7%; Score 1067; DB 12; Length 297;  
 Best Local Similarity 88.7%; Pred. No. 2.1e-76;  
 Matches 205; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSNL 60  
 DB 67 HKVTSITIIQDATTNQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSTL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 QSTVGTGKNTTTTQAPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTTIT 180  
 DB 187 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTIIT 246  
 QY 181 LLTNNTTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 231  
 DB 247 LLTSNTTRNPELTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 297

## RESULT 8

Q82066 ID Q82066 PRELIMINARY; PRT; 297 AA.  
 AC Q82066;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glycoprotein.

GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI\_TaxID=11250;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Subgroup A;

RX MEDLINE=94335057; PubMed=8057427;

RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,

RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,

RA Melero J.A.;

RT "Evolutionary pattern of human respiratory syncytial virus (subgroup

RT A): cocirculating lineages and correlation of genetic and antigenic

RT changes in the G glycoprotein.";

RL J. Virol. 68:5448-5459 (1994).

DR EMBL; Z33427; CAA83870.1; -;

DR InterPro; IPR000925; Glycoprot G.

DR InterPro; IPR006162; Ppantne attach.

DR Pfam; PF00802; Glycoprotein G; 1.

DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.

SQ SEQUENCE 297 AA; 32589 MW; F210F8C68193F5DD CRC64;

Query Match 86.7%; Score 1066; DB 12; Length 297;  
 Best Local Similarity 88.3%; Pred. No. 2.5e-76;  
 Matches 204; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSNL 60  
 DB 67 HKVTSITIIQDATTNQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSTL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 120

DB 127 QSTVGTGKNTTTTQAPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTIIT 180  
 DB 187 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTIIT 246  
 QY 181 LLTNNTTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 231  
 DB 247 LLTSNTTRNPELTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTSR 297

## RESULT 9

Q9YVB3 ID Q9YVB3 PRELIMINARY; PRT; 293 AA.

AC Q9YVB3;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Attachment glycoprotein G (Fragment).

OS Human respiratory syncytial virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI\_TaxID=11250;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MV6973;

RX MEDLINE=99022964; PubMed=9806017;

RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;

RT "Antigenic and genetic diversity among the attachment proteins of

RT group A respiratory syncytial viruses that have caused repeat

RT infections in children.";

RL J. Infect. Dis. 178:925-932 (1998).

DR EMBL; AF065407; AAD02943.1; -;

DR InterPro; IPR000925; Glycoprot G.

DR InterPro; IPR006162; Ppantne attach.

DR Pfam; PF00802; Glycoprotein G; 1.

DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.

FT NON\_TER 1

SQ SEQUENCE 293 AA; 32125 MW; 17B5B43396A63CCF CRC64;

Query Match 86.5%; Score 1063; DB 12; Length 293;  
 Best Local Similarity 88.8%; Pred. No. 4.3e-76;  
 Matches 206; Conservative 6; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDATSQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSNL 60  
 DB 62 HKVTSITIIQDATSQIKNTTPTLTQDQGLGFSNLSSEITSTTILASTTPGVKSTL 121  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 122 QSTVGTGKNTTTTQAPSKPTTKQKPKPKPKNDHFEVFNVPVCSICSNPTCWAIC 181  
 QY 121 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTIIT 180  
 DB 182 KRIPNKKPGKKTTPKTKPKTKTKKDLKPQTKPKKEVPTTKTEPTTNTKNTIIT 241  
 QY 181 LLTNNTTGNPKLTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTTR 232  
 DB 242 LLTSNTTRNPELTSOMETFTSTSEGNLSPSQVSTTSEHPSQSPSPNTPR 293

## RESULT 10

Q9YVB4 ID Q9YVB4 PRELIMINARY; PRT; 298 AA.

AC Q9YVB4;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Attachment glycoprotein G.

OS Human respiratory syncytial virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI\_TaxID=11250;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=WV5222;  
 RA MEDLINE=99022964; PubMed=9806017;  
 RX Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;  
 "Antigenic and genetic diversity among the attachment proteins of  
 RT group A respiratory syncytial viruses that have caused repeat  
 RT infections in children";  
 RL J. Infect. Dis. 178:925-932(1998).  
 DR EMBL; AF065406; AAD02942.1; --  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.  
 SQ SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 86.4%; Score 1062; DB 12; Length 298;  
 Best Local Similarity 87.5%; Pred. No. 5.2e-76;  
 Matches 203; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 126  
 QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 180  
 DB 187 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 246  
 QY 181 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 231  
 DB 247 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 297

RESULT 11  
 Q82071 ID Q82071 PRELIMINARY; PRT; 297 AA.  
 AC Q82071;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glycoprotein.  
 GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=subgroup A;  
 RX MEDLINE=94335057; PubMed=8057427;  
 RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrasile S., Russi J.,  
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,  
 RA Melero J.A.;  
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup  
 RT A): cocirculating lineages and correlation of genetic and antigenic  
 RT changes in the G glycoprotein";  
 RL J. Virol. 68:5448-5459(1994).  
 DR EMBL; Z33432; CA83875.1; --  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.  
 SQ SEQUENCE 297 AA; 32581 MW; AD947B2318193306 CRC64;

Query Match 85.9%; Score 1056; DB 12; Length 297;  
 Best Local Similarity 87.4%; Pred. No. 1.5e-75;  
 Matches 202; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 126  
 QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 180  
 DB 187 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 246  
 QY 181 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 231  
 DB 247 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 297

RESULT 12  
 Q906T3 ID Q906T3 PRELIMINARY; PRT; 278 AA.  
 AC Q906T3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glycoprotein (Fragment).  
 GN G.  
 OS Human respiratory syncytial virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11250;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=93006;  
 RX MEDLINE=20283719; PubMed=10823752;  
 RA Choi B.H., Lee H.J.;  
 RT "Genetic diversity and molecular epidemiology of the G protein of  
 RT subgroups A and B of respiratory syncytial virus isolated over 9  
 RT consecutive epidemics in Korea";  
 RL J. Infect. Dis. 181:1547-1556(2000).  
 DR EMBL; AF193313; AAF23736.1; --  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.  
 FT NON TER 1  
 SQ SEQUENCE 278 AA; 30443 MW; 8E2974F2E1AE15C9 CRC64;

Query Match 85.8%; Score 1055; DB 12; Length 278;  
 Best Local Similarity 87.4%; Pred. No. 1.7e-75;  
 Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 DB 48 HKVTLTALIQDTSQIKNTPTYLTPQDQPGISFNSLSEITSTQTTILASTTPGVKSNL 107  
 QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 108 QSTTVGKTNTTTTQTPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 167  
 QY 121 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 180  
 DB 168 KRIPNKKPGKTTTKPTTKPTTKKDKLPQTTKPKVPTTKTEPTINTTKNITTT 227  
 QY 181 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 230  
 DB 228 LTNNTGNPKLTSMQETHSTSSGNLSPSQVSTTSBHPSPSPPTTR 277

RESULT 13  
 Q9YVB2 ID Q9YVB2 PRELIMINARY; PRT; 292 AA.  
 AC Q9YVB2;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)

```

DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Attachment glycoprotein G (Fragment).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WV19983;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Murson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065408; AA02944.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein_G; 1.
DR PRINTS; PR01217; PRICHEXTENSH.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
FT NON_TER
SQ SEQUENCE 292 AA; 31964 MW; 8942A8DD0A402A4B CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 292;
Best Local Similarity 87.9%; Pred. No. 2.6e-75;
Matches 203; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSNLSLSEITSTQTTILASTTPGVKSNL 60
DB 62 HKVLTSTIIQDAINQIKNTPTVLTQNPQLGISPNLSITSLIILOSTTPGVKSTL 121
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 120
DB 122 QSTTVGKNTTTTQAPNKPETTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 181
QY 121 KRIPNKKPGKTTTKPKTKPTKTKDKLPQTKKEVPTTKTEPTINTKTNIITT 180
DB 182 KRIPNKKPGKTTTKPKTKPTTKTKGKPKPTTKSKEVPTTKTEPTINTKTNIITT 241
QY 181 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 231
DB 242 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 292

RESULT 14
Q82058 PRELIMINARY; PRT; 297 AA.
AC Q82058;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE (MAD-1-89) subgroup A, G glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=subgroup A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL; Z33456; CAA83879.1; -.
DR InterPro; IPR000925; Glycoprot_G.

```

```

DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein_G; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
SQ SEQUENCE 297 AA; 32741 MW; 0E567A174BF64964 CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 297;
Best Local Similarity 87.4%; Pred. No. 2.7e-75;
Matches 201; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSNLSLSEITSTQTTILASTTPGVKSNL 60
DB 67 HKVLTALIIQDATSQIKNTPTVLTQNPQLGISPNLSITSTQTTILASTTSAESTP 126
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 120
DB 127 QSTTVKNTTTTQIQPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 186
QY 121 KRIPNKKPGKTTTKPKTKPTKTKDKLPQTKKEVPTTKTEPTINTKTNIITT 180
DB 187 KRIPNKKPGKTTTKPKTKPTTKTKDKLPQTKKEVPTTKTEPTINTKTNIITT 246
QY 181 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 230
DB 247 LLTNTTGNPKLTSQMETFHTSSEGNLSPSQVSTTSEHPSQSPSPNTT 296

RESULT 15
Q86359 PRELIMINARY; PRT; 295 AA.
AC Q86359;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE G protein (fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73353; CAA51764.1; -.
DR InterPro; IPR000925; Glycoprot G.
DR InterPro; IPR006162; Ppantne attach.
DR Pfam; PF00802; Glycoprotein_G; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
FT NON_TER
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match 85.6%; Score 1052; DB 12; Length 295;
Best Local Similarity 87.4%; Pred. No. 3.2e-75;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTALIIQDATSQIKNTPTVLTQDPQLGIFSNLSLSEITSTQTTILASTTPGVKSNL 60
DB 65 HKVLTALIIQDATSQIKNTPTVLTQNPQLGISPNLSITSTQTTTPAPTSAESTP 124
QY 61 OPTTVKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 120
DB 125 QSTTVKNTTTTQIQPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNNPTCWAIC 184
QY 121 KRIPNKKPGKTTTKPKTKPTKTKDKLPQTKKEVPTTKTEPTINTKTNIITT 180

```

Db 185 KRIENKKGKTTTKPTTKPTIKTTKKDLKPQTTKPKVLTTRKTEKPTINTTRTNIRTT 244  
Qy 181 LLTNTTGNPKLTSQMETFHSSTSEGNLSPSQVSTTSEHPSQPSPSPNTT 230  
Db 245 LLTNTTGNPEYTSQKETLHSTSEGNPSQVYTTSEYPSQPSPSPNTT 294

Search completed: October 29, 2003, 17:40:53  
Job time : 55.717 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:12:25 ; Search time 12.6943 Seconds  
(without alignments)  
859.454 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTTHAIQDASQIKNT.....VSTTSEHPSQSPSPNTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1216	98.9	298	1	VGLG_HRSVL
2	1120	91.1	298	1	VGLG_HRSVA
3	1052	85.6	297	1	VGLG_HRSV6
4	1045	85.0	297	1	VGLG_HRSV3
5	1037	84.4	297	1	VGLG_HRSV7
6	1033	84.1	297	1	VGLG_HRSV4
7	1019	82.9	298	1	VGLG_HRSV5
8	1009	82.1	297	1	VGLG_HRSV2
9	505.5	41.1	292	1	VGLG_HRSV8
10	497	40.4	292	1	VGLG_HRSV1
11	202	16.4	307	1	SGS3_DROME
12	198.5	15.2	2812	1	ZAN_HUMAN
13	193	15.7	5179	1	MUC2_HUMAN
14	191	15.5	662	1	MUC1_XENLA
15	183	14.9	217	1	SGS3_DROSI
16	183	14.9	1161	1	DANA_YEAST
17	180.5	14.7	263	1	VGLG_BRSTV1
18	179.5	14.6	263	1	VGLG_BRSTV4
19	177	14.4	257	1	VGLG_BRSTV3
20	176.5	14.4	263	1	VGLG_BRSTV6
21	170.5	13.9	263	1	VGLG_BRSTV2
22	169.5	13.8	1367	1	AMTH_YEAST
23	169	13.8	257	1	ZAN_RABIT
24	166	13.5	2282	1	ZAN_PIG
25	164.5	13.4	2476	1	ZAN_PIG
26	160	13.0	263	1	SGS3_DROVA
27	159.5	13.0	263	1	VGLG_ORSVW
28	159	12.9	257	1	VGLG_BRSTV5
29	156	12.7	257	1	VGLG_BRSTV1
30	154	12.5	257	1	VGLG_BRSTV2
31	154	12.5	1233	1	MUSA_HUMAN
32	153.5	12.5	5703	1	MUSEB_HUMAN
33	150.5	12.2	1513	1	MUC2_RAT

## RESULT 1

VGLG\_HRSVL  
ID VGLG\_HRSVL STANDARD; PRT; 298 AA.  
AC P20895;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Major surface glycoprotein G (Attachment glycoprotein G).  
GN G.  
OS Human respiratory syncytial virus (subgroup A / strain Long).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
OX NCBI\_TaxID=11260;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8728957; PubMed=2441388;  
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
RT "The G glycoprotein of human respiratory syncytial viruses of  
RT subgroups A and B: extensive sequence divergence between  
RT antigenically related proteins";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).  
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
CC HEMAGGLUTININATING ACTIVITIES.  
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M17212; AAA47411.1; -  
PIR; A32703; MGNZRL.  
InterPro; IPR000925; Glycoprot G.  
Pfam; PF00802; Glycoprotein G; 1.  
Transmembrane; Glycoprotein.  
DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 66 POTENTIAL.  
FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 298 AA; 32781 MW; B79FEFA4BA73B0E CRC64;

Query Match 98.9%; Score 1216; DB 1; Length 298;  
Best Local Similarity 99.1%; Pred. No. 4.3e-75;  
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M1486; AAB59857.1; -  
 DR EMBL; X03149; CAA26928.1; -  
 DR EMBL; U50362; AAB86663.1; -  
 DR EMBL; U50363; AAB86675.1; -  
 DR EMBL; U63644; AAC55969.1; -  
 DR PIR; A94048; MGZ;  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 298 AA; 32586 MW; 993C3D2D68BC634 CRC64;  
 Query Match 91.1%; Score 1120; DB 1; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.2e-68;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 126  
 QY 61 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHVFVPCISNNPTCWAIC 120  
 DB 127 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHVFVPCISNNPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKTKKOLKPKQTTKPEVPTTKPTERTPTNTTKNITTT 180  
 DB 187 KRIPNKKPKGKTTTKPTKPTKTKTKKOLKPKQTTKPEVPTTKPTERTPTNTTKNITTT 246  
 QY 181 LLTNNTTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 DB 247 LLTNNTTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 RESULT 2  
 VGLG\_HRSVA STANDARD; PRT; 298 AA.  
 AC P03423;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain A2).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11259;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=85216636; PubMed=3858865;  
 RA Wertz G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;  
 RT "Nucleotide sequence of the G protein gene of human respiratory  
 RT syncytial virus reveals an unusual type of viral membrane protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=86067198; PubMed=4069997;  
 RA Satake M., Colligan J.B., Elango N., Norrby E., Venkatesan S.;  
 RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel  
 RT structure.";  
 RL Nucleic Acids Res. 13:7795-7812(1985).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95266253; PubMed=7747420;  
 RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;  
 RT "A cold-passaged, attenuated strain of human respiratory syncytial  
 RT virus contains mutations in the F and L genes.";  
 RL Virology 208:478-484(1995).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=97187925; PubMed=9035372;  
 RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,  
 RA Murphy B.R.;  
 RT "Acquisition of the ts phenotype by a chemically mutagenized cold-  
 RT passaged human respiratory syncytial virus vaccine candidate results  
 RT from the acquisition of a single mutation in the polymerase (L)  
 RT gene.";  
 RL Virus Genes 13:269-273(1996).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M1486; AAB59857.1; -  
 DR EMBL; X03149; CAA26928.1; -  
 DR EMBL; U50362; AAB86663.1; -  
 DR EMBL; U50363; AAB86675.1; -  
 DR EMBL; U63644; AAC55969.1; -  
 DR PIR; A94048; MGZ;  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66  
 FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 298 AA; 32586 MW; 993C3D2D68BC634 CRC64;  
 Query Match 91.1%; Score 1120; DB 1; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.2e-68;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSEITSTQTTILASTTPGVKSNL 126  
 QY 61 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHVFVPCISNNPTCWAIC 120  
 DB 127 OPTTVKTKNTTTTQPSKPTTKQKQKPNKPNNDFFHVFVPCISNNPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKTKKOLKPKQTTKPEVPTTKPTERTPTNTTKNITTT 180  
 DB 187 KRIPNKKPKGKTTTKPTKPTKTKTKKOLKPKQTTKPEVPTTKPTERTPTNTTKNITTT 246  
 QY 181 LLTNNTTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 DB 247 LLTNNTTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 RESULT 3  
 VGLG\_HRSV6 STANDARD; PRT; 297 AA.  
 AC P27025;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain r6256).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11256;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 RT subgroup A respiratory syncytial viruses.";  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC -----  
 CC PIR; J01208; J01208.  
 DR InterPro; IPR000925; Glycoprot G.  
 DR Pfam; PF00802; Glycoprotein\_G; 1.



Transmembrane; Glycoprotein.  
 KW DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 103 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;  
 Query Match 85.6%; Score 1052; DB 1; Length 297;  
 Best Local Similarity 87.4%; Pred. No. 4.2e-64;  
 Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 120  
 DB 127 QSTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 180  
 DB 187 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 246  
 QY 181 LITNNTGNPKLTQMETFHSSTSGNLSPSQVSTTSEHPSPQSPSPNTT 230  
 DB 247 LITNNTGNPKLTQMETFHSSTSGNLSPSQVSTTSEHPSPQSPSPNTT 297  
 RESULT 4  
 VGLG\_HRSV3  
 ID VGLG\_HRSV3 STANDARD; PRT; 297 AA.  
 AC P27022;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain rab1734).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11253;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 subgroup A respiratory syncytial viruses.";  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 PIR; JQ1205; JQ1205.  
 DR InterPro; IPR000925; Glycoprotein G.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 297 AA; 32525 MW; 48448F9E091E1802 CRC64;  
 Query Match 85.0%; Score 1045; DB 1; Length 297;  
 Best Local Similarity 86.6%; Pred. No. 1.3e-63;  
 Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 120  
 DB 127 QSTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 180

Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 120  
 DB 127 QSTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 180  
 DB 187 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 246  
 QY 181 LITNNTGNPKLTQMETFHSSTSGNLSPSQVSTTSEHPSPQSPSPNTT 231  
 DB 247 LITNNTGNPKLTQMETFHSSTSGNLSPSQVSTTSEHPSPQSPSPNTT 297  
 RESULT 5  
 VGLG\_HRSV7  
 ID VGLG\_HRSV7 STANDARD; PRT; 297 AA.  
 AC P27026;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (strain rab6614).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11257;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 subgroup A respiratory syncytial viruses.";  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 PIR; JQ1209; JQ1209.  
 DR InterPro; IPR000925; Glycoprotein G.  
 DR Pfam; PF00802; Glycoprotein G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 297 AA; 32670 MW; 58B384028E437ACD CRC64;  
 Query Match 84.4%; Score 1037; DB 1; Length 297;  
 Best Local Similarity 86.5%; Pred. No. 4.3e-63;  
 Matches 199; Conservative 5; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDQTSQIKNTTPTLTQDPQLGIFSLSNLSSETTSQTTLASTTPGVKSNL 126  
 QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 120  
 DB 127 QSTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKPTKPTKTKKQKQKPNKPNNDHFEVFNFPVCSICSNNTPTCWAIC 180



RN SEQUENCE FROM N.A.  
 RP MEDLINE=91374005; PubMed=1895054;  
 RA Cane P.A., Matthews D.A., Pringle C.R.;  
 RT "Identification of variable domains of the attachment (G) protein of  
 RT subgroup A respiratory syncytial viruses.";  
 RL J. Gen. Virol. 72:2091-2096(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; M55633; AAA47413.1; -;  
 CC EMBL; M73545; AAA47408.1; -;  
 CC PIR; A37077; MGN260.  
 CC InterPro: IPR000925; Glycoprotein G.  
 CC Pfam; PF00802; Glycoprotein G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 297 AA; 32745 MW; FC72A7F3A8EF67C CRC64;  
  
 Query Match 82.1%; Score 1009; DB 1; Length 297;  
 Best Local Similarity 84.0%; Pred. No. 3.2e-61;  
 Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;  
  
 QY 1 HKVLTITAIIDATNSQIKNTTPTLTQDPQIGISFNSLSEITSTTTILASTTGVKSNL 60  
 DB 67 HRVSTSTIIQDATNQIKNTTPTLTQDPQIGISFNSLSEITSTTTILASTTGVKSNL 126  
  
 QY 61 QPTTVKTKNTTTTQPSKPTTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 120  
 DB 127 QSTTVIKNTTTTQAPKSKTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 186  
  
 QY 121 KRIPNKKPGKTTTKPTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 180  
 DB 187 KRIPNKKPGKTTTKPTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 246  
  
 QY 181 LLTNTNTGNPKLTQSMETFTSTSEGNLSPSOVSTTSEHPSQSPPTNR 231  
 DB 247 PLTNTARNPLTQSMETFTSTSEGNLSPSOVSTTSEHPSQSPPTNR 297  
  
 RESULT 9  
 VGLG HRSV8 STANDARD; PRT; 292 AA.  
 AC P23041.  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 8/60).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11258;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90357765; PubMed=1697126;  
 RA Sullender W.M., Anderson K., Wertz G.W.;  
 RT "The respiratory syncytial virus subgroup B attachment glycoprotein:  
 RT analysis of sequence, expression from a recombinant vector, and  
 RT evaluation as an immunogen against homologous and heterologous  
 RT subgroup virus challenge.";  
 RL Virology 178:195-203(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91374595; PubMed=1895391;  
 RA Sullender W.M., Mufson M.M., Anderson L.J., Wertz G.W.;

RT "Genetic diversity of the attachment protein of subgroup B  
 RT respiratory syncytial viruses.";  
 RL J. Virol. 65:5425-5434(1991).  
 CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE  
 CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND  
 CC HEMAGGLUTININATING ACTIVITIES.  
 CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED  
 CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.  
 CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS  
 CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; M55633; AAA47413.1; -;  
 CC EMBL; M73545; AAA47408.1; -;  
 CC PIR; A37077; MGN260.  
 CC InterPro: IPR000925; Glycoprotein G.  
 CC Pfam; PF00802; Glycoprotein G; 1.  
 KW Transmembrane; Glycoprotein.  
 FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 38 66 POTENTIAL.  
 FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 292 AA; 32143 MW; 8EC60C85EF057BBS CRC64;  
  
 Query Match 41.1%; Score 505.5; DB 1; Length 292;  
 Best Local Similarity 47.6%; Pred. No. 1.9e-27;  
 Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;  
  
 QY 1 HKVLTITAIIDATNSQIKNTTPTLTQDPQIGISFNSLSEITSTTTILASTTGVKSNL 60  
 DB 67 HKVLTITVTVQPIKNTTKNTSTLTQVPPERVNSKQPTTSTPIHNSATISPTKSET 126  
  
 QY 61 QPTTVKTKNTTTTQPSKPTTKQKPNKPNNDHFVFNVPVCSICSNPTCWAIC 120  
 DB 127 HHTTAQTKGRITTSQTNKPSKSKNPKPKDDHFFVFNVPVCSICGNQLCKSIC 186  
  
 QY 121 KRIPNKKPGKTTTKPTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 179  
 DB 187 KRIPNKKPGKTTTKPTKPKTKTKKDLKPKQTKPKKEVPTTKPTTEPTINTKNTTT 246  
  
 QY 180 TLLTNTNTGNPKLTQSMETFTSTSEGNLSPSOVSTTSEHPSQSPSP 226  
 DB 247 SOSTVIDITPKYTIQQOSSLHSTTSSTPSSQTPIPTASE-PS-TSNP 291  
  
 RESULT 10  
 VGLG HRSV1 STANDARD; PRT; 292 AA.  
 AC P20896;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DE Major surface glycoprotein G (Attachment glycoprotein G).  
 GN G.  
 OS Human respiratory syncytial virus (subgroup B / strain 18537).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.  
 OX NCBI\_TaxID=11251;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=87289657; PubMed=2441388;  
 RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;  
 RT "The G glycoprotein of human respiratory syncytial viruses of  
 RT subgroups A and B: extensive sequence divergence between

antigenically related proteins.";

Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629 (1987).

-1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE HEMAGGLUTININATING ACTIVITIES.

-1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.

-1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M17213; AAA47412.1; --

PIR; B32703; MGZ18.

InterPro; IPR000925; Glycoprot G.

Pfam; PF00802; Glycoprotein\_G; 1.

Transmembrane; Glycoprotein.

DOMAIN 1 37

TRANSMEM 38 66

DOMAIN 67 292

EXTRACELLULAR (POTENTIAL).

CARBOHYD 81 81

N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 86 86

N-LINKED (GLCNAC. . .) (POTENTIAL).

CARBOHYD 100 100

N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 292 AA; 32306 MW; BC8C59F69CA7AFC2 CRC64;

Query Match 40.48; Score 497; DB 1; Length 292;

Best Local Similarity 46.8; Pred. No. 6.9e-27;

Matches 104; Conservative 28; Mismatches 98; Indels 2; Gaps 2;

QY 1 HKVLTALIQDAPSOIKNTPTTYLTQDQGLGIFSNLSITSTTILASTTGVKSNL 60

DB 67 HKVLTITVTQTIKNHTEKNIISTVLTQVPERVNSKQPTTSPHNSATISPTWSET 126

QY 61 QPTVTKNTTITQTSKPTTKORQKNPKNDHFVFNVPVCSISNPTCAIC 120

DB 127 HHTAQTGKRITSTOTNKESTKRSRKNPKPKDDYHFEVFNVPVCSICGNOLCKSK 186

QY 121 KRIPNKPKGKTTTKPKTKPTTKTT-KDKLKPQTKPEVPTTKPTTPTTNTKNTT 179

DB 187 KTIPTNPKPKPKTKPKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 246

QY 180 TLTNTNTGNPKLTSQMTHTSTSTSEGNLSPSQVSTTSEHPS 221

DB 247 SQSTVLDTITPKVTIQOQLSHSTSTSTPTSPSTQIPTASE-PS 287

RESULT 11

SGS3 DROME STANDARD; PRT; 307 AA.

AC P02840; Q9VJT2;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Salivary glue protein Sgs-3 precursor.

GN Sgs3 OR CG11720.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxId=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83294545; PubMed=6411930;

RA Garfinkel M.D., Pruitt R.E., Meyerowitz E.M.;

RT "DNA sequences, gene regulation and modular protein evolution in the Drosophila 68C glue gene cluster.";

J. Mol. Biol. 168:765-789 (1983).

RN RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Morkulov G., Malyshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Poirard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195 (2000).

[3]

RN RP SEQUENCE OF 1-28 FROM N.A.

RX MEDLINE=88332966; PubMed=3138416;

RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;

RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";

J. Mol. Biol. 201:273-287 (1988).

[4]

RN RP DEVELOPMENTAL STAGE.

RX MEDLINE=94038699; PubMed=8223281;

RA Huet F., Ruiz C., Richards G.;

RT "Puffs and PCR: the in vivo dynamics of early gene expression during ecdysonic responses in Drosophila.";

Development 118:613-627 (1993).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: SALIVARY GLAND SPECIFIC.

CC -1- DEVELOPMENTAL STAGE: IN THE SALIVARY GLANDS OF MID INSTAR LARVAE LEVELS DRAMATICALLY INCREASE DURING PUFF STAGE 1 AT 98-106 HOURS OF DEVELOPMENT. LEVELS REMAIN CONSTANT AND ABUNDANT IN LATE LARVAE UNTIL PUFF STAGE 10, THEN DECREASE BY STAGE 11.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X01918; CAA25994.1; --

DR EMBL; AE003544; BAF50056.1; --

DR EMBL; X78392; CAA55154.1; --

DR PIR; A03329; GSPF3.  
DR FlyBase; FBgn0003373; Sgs3.  
KW Repeat; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 307 SALIVARY GLUE PROTEIN SGS-3.  
SQ SEQUENCE 307 AA; 32196 MW; 45803DED6C418BC CRC64;  
  
Query Match 16.4%; Score 202; DB 1; Length 307;  
Best Local Similarity 32.8%; Pred. No. 4.4e-07;  
Matches 77; Conservative 17; Mismatches 93; Indels 48; Gaps 13;  
  
QY 4 TLTTALIQDASQIKNTPTVLTQDQGLGISFSLSEITSTQTTILASTPGV--KSNLQ 61  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : :  
49 TITTTTCAPPQQ--STQPPCTTSKP-----TTPKQTTTQLPCTTPTTTTKATTK 97  
QY 62 PTTVKTKNTTQTQPSKPTTKQKPPKPNNDFFHVFVPCISNNPTCWAICK 121  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : :  
98 PTTTKATTKATTKPT--TTKQTTQLP-----CITPTTKQTTT 136  
QY 122 RIPNKKP--GKTTTKP--TKPTTKTK-KDLPQTKPKVEPTTKP--TSEPTINTTKTN 176  
Db : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : :  
137 QLPCTPTPTTKPTTKPTTKPTTKPT--TTKPTTKPTTKPT--TTKPTTKPT--TTKPT 189  
QY 177 ITTLLTNTTGNPKLSQMTFSTSEGNLSQVSTTSEHPSQSPSPPTTR 231  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : :  
190 TTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTK 242  
  
RESULT 12  
ZAN\_HUMAN STANDARD; PRT; 2812 AA.  
ID ZAN\_HUMAN AC Q9Y493; Q00218; Q96L85; Q96L86; Q96L87; Q96L88; Q96L89; Q96L90;  
AC Q9XN9; Q9BZ83; Q9BZ84; Q9BZ85; Q9BZ86; Q9BZ87; Q9BZ88;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Zonadhesin precursor.  
GN ZAN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).  
RC TISSUE-Testis;  
RA Chung T.L., Wasler M.J., Cornwall G.A., Hardy D.M.;  
RT "Multiple intra-species variants of human zonadhesin."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [2] PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE=99018118; PubMed=9799793;  
RA Tsui L.-C., Rosenthal A.;  
RT "Large-scale sequencing of two regions in human chromosome 7q22:  
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci  
RT reveals 17 genes."  
RL Genome Res. 8:1060-1073(1998).  
RN [3] SEQUENCE OF 1810-2812 FROM N.A. (ISOFORM 1).  
RX MEDLINE=21138439; PubMed=11239002;  
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,  
RA Chung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,  
RA Miller W., Koop B.F.;  
RT "Comparative analysis of the gene-dense ACHE/TFR2 region on human  
RT chromosome 7q22 with the orthologous region on mouse chromosome 5";  
RL Nucleic Acids Res. 29:1352-1365(2001).  
RN [4] SEQUENCE OF 2375-2683 FROM N.A. (ISOFORM 7).  
RX TISSUE-Testis;  
RA MEDLINE=97271566; PubMed=9126492;  
RA Gao Z., Harumi T., Garbers D.L.;  
RT "Chromosome localization of the mouse zonadhesin gene and the human  
RT zonadhesin gene (ZAN).";

Genomics 41:119-122 (1997).  
-!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA  
OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR  
SIGNALING.  
-!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.  
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE  
APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).  
-!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=7;  
Name=3;  
IsoId=Q9Y493-1; Sequence=Displayed;  
Name=1;  
IsoId=Q9Y493-2; Sequence=VSP\_001430, VSP\_001431;  
Name=2;  
IsoId=Q9Y493-3; Sequence=VSP\_001428, VSP\_001429;  
Name=4;  
IsoId=Q9Y493-4; Sequence=VSP\_001424, VSP\_001425;  
Name=5;  
IsoId=Q9Y493-5; Sequence=VSP\_001420, VSP\_001421;  
Name=6;  
IsoId=Q9Y493-6; Sequence=VSP\_001422, VSP\_001423;  
Name=7;  
IsoId=Q9Y493-7; Sequence=VSP\_001426, VSP\_001427;  
-!- TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.  
-!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE  
ZONA PELLUCIDA.  
-!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,  
THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF  
SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.  
-!- DOMAIN: The VWFD domain 2 may mediate covalent oligomerization (by  
similarity to human intestinal mucin MUC2).  
-!- SIMILARITY: Contains 3 MAM domains.  
-!- SIMILARITY: Contains 4 VWFD domains.  
-!- SIMILARITY: Contains 1 EGF-like domain.  
-!- CAUTION: Ref.2 sequence differs from that shown due to  
transposition of a number of exons.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF332975; AAK01431.1; -  
EMBL; AF332976; AAK01432.1; -  
EMBL; AF332977; AAK01433.1; -  
EMBL; AF332978; AAK01434.1; -  
EMBL; AF332979; AAK01435.1; -  
EMBL; AF332980; AAK01436.1; -  
EMBL; AY046055; AAL04410.1; -  
EMBL; AY046055; AAL04411.1; -  
EMBL; AY046055; AAL04412.1; -  
EMBL; AY046055; AAL04413.1; -  
EMBL; AY046055; AAL04414.1; -  
EMBL; AY046055; AAL04415.1; -  
EMBL; AF053356; AAC78790.1; ALT\_SEQ.  
EMBL; AF312032; AAK21011.1; -  
EMBL; U83191; AAC51208.1; -  
Genew; HGNC:12857; ZAN.  
MIM; 602372; -  
GO; GO:0016324; C:apical plasma membrane; NAS.  
GO; GO:0007339; P:binding of sperm to zona pellucida; NAS.  
GO; GO:0008037; P:cell recognition; NAS.  
GO; GO:0016337; P:cell-cell adhesion; NAS.  
InterPro; IPR006209; EGF like.  
InterPro; IPR000998; MAM domain.  
InterPro; IPR002919; TIL Cysrich.  
InterPro; IPR003328; TILa Cysrich.  
InterPro; IPR001007; VWF\_C.  
InterPro; IPR001846; VWF\_D.  
Pfam; PF00629; MAM; 3.







[illegible]







C:Accession: A94048; A93599; A04039  
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
A:Reference number: A94048; MUID:85216636; PMID:3858865  
A:Accession: A94048  
A:Molecule type: mRNA  
A:Residues: 1-298 <WER>  
A:Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:K03349; GB:M11217; GB:7.1; PID:G333932  
A:Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
A:Note: this protein may carry 40-80 separate O-linked carbohydrate chains distributed a  
R:Satake, M.; Coligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S.  
Nucleic Acids Res. 13, 7795-7812, 1985  
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
A:Reference number: A93599; MUID:86067198; PMID:4069997  
A:Accession: A93599  
A:Molecule type: mRNA  
A:Residues: 1-298 <SAT>  
A:Cross-references: GB:X03149; NID:G60997; PIDN:CAA26928.1; PID:G60998  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:38-66/Domain: transmembrane #status predicted <TM>  
F:85,103,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 91.1%; Score 1120; DB 1; Length 298;  
Best Local Similarity 92.7%; Pred. No. 1.6e-68;  
Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
QY 1 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 246  
QY 181 LITNNTTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 230  
DB 247 LLTSTNTRNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 296  
RESULT 3  
JQ1208  
attachment protein - human respiratory syncytial virus (strain RSB6256)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1208  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2031-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1208  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 85.0%; Score 1045; DB 2; Length 297;  
Best Local Similarity 86.6%; Pred. No. 1.8e-63;  
Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;  
QY 1 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 246  
QY 181 LITNNTTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 231  
DB 247 LLTSTNTRNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 297  
RESULT 5  
JCS580  
G protein - Human respiratory syncytial virus  
C:Species: Human respiratory syncytial virus  
C>Date: 11-Nov-1997 #sequence\_revision 11-Nov-1997 #text\_change 26-Feb-1998  
C:Accession: JCS580  
R:Geng, X.; Wang, Z.; Qian, X.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.  
Chinese J. Virol. 12, 317-322, 1996  
A:Title: Molecular analysis of G protein gene of a respiratory syncytial virus strain is  
A:Reference number: JCS580  
A:Accession: JCS580  
A:Molecule type: mRNA  
A:Residues: 1-298 <GEN>  
A:Experimental source: strain B79  
A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
F:1-38/Domain: intracellular #status predicted <INT>  
F:39-66/Domain: transmembrane #status predicted <TM>

C:Accession: A94048; A93599; A04039  
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985  
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus  
A:Reference number: A94048; MUID:85216636; PMID:3858865  
A:Accession: A94048  
A:Molecule type: mRNA  
A:Residues: 1-298 <WER>  
A:Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:K03349; GB:M11217; GB:7.1; PID:G333932  
A:Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein  
A:Note: this protein may carry 40-80 separate O-linked carbohydrate chains distributed a  
R:Satake, M.; Coligan, J.E.; Elango, N.; Norrby, E.; Venkatesan, S.  
Nucleic Acids Res. 13, 7795-7812, 1985  
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.  
A:Reference number: A93599; MUID:86067198; PMID:4069997  
A:Accession: A93599  
A:Molecule type: mRNA  
A:Residues: 1-298 <SAT>  
A:Cross-references: GB:X03149; NID:G60997; PIDN:CAA26928.1; PID:G60998  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:38-66/Domain: transmembrane #status predicted <TM>  
F:85,103,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 91.1%; Score 1120; DB 1; Length 298;  
Best Local Similarity 92.7%; Pred. No. 1.6e-68;  
Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
QY 1 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSTL 126  
QY 61 QPTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
DB 127 QSTTVKTKNTTTQTPSKPTTKQKQKPNKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 180  
DB 187 KRIPNKKPGKTTTKPTKKPTFKTTKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 246  
QY 181 LITNNTTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 232  
DB 247 LLTSTNTRNPELTSMQETFHSTSEGNLSPSQVSTTSEHPSPQSPSPNTT 298  
RESULT 3  
JQ1208  
attachment protein - human respiratory syncytial virus (strain RSB6256)  
N:Alternate names: G protein  
C:Species: human respiratory syncytial virus  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
C:Accession: JQ1208  
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
J. Gen. Virol. 72, 2031-2096, 1991  
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A  
A:Reference number: JQ1204; MUID:91374005; PMID:1895054  
A:Accession: JQ1208  
A:Molecule type: mRNA  
A:Residues: 1-297 <CAN>  
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract il  
children and adults.  
C:Superfamily: respiratory syncytial virus major surface glycoprotein G  
C:Keywords: glycoprotein; transmembrane protein  
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predic  
Query Match 85.6%; Score 1052; DB 2; Length 297;  
Best Local Similarity 87.4%; Pred. No. 6.2e-64;  
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;  
QY 1 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSNL 60  
DB 67 HKVLTLLAIIDATSOIKNTTPTLTQDPQLGIFSNLSITSTTTILASTTPGVKSTL 126



RESULT 9  
 JQ1204  
 Attachment protein - human respiratory syncytial virus (strain RSB642)  
 N/Alternate names: G protein  
 C/Species: human respiratory syncytial virus  
 C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Feb-1997  
 C/Accession: JQ1204  
 R/Cane, P.A.; Matthews, D.A.; Pringle, C.R.  
 J. Gen. Virol. 72, 2091-2096, 1991  
 A/Title: Identification of variable domains of the attachment (G) protein of subgroup A  
 A/Reference number: JQ1204; MUID:91374005; PMID:1895054  
 A/Accession: JQ1204  
 A/Molecule type: mRNA  
 A/Residues: 1-297 <CAN>  
 C/Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill  
 children and adults.  
 C/Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C/Keywords: glycoprotein; transmembrane protein  
 F:95,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 82.1%; Score 1009; DB 2; Length 297;  
 Best Local Similarity 84.0%; Pred. No. 4.8e-61;  
 Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 HKVLTLLAIQDANSQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKSNL 60  
 DB 67 HRVTSITIIQDANQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKLT 126  
 QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPNPNNDPHEFVFNVPVCSICSNNPCTCAIC 120  
 DB 127 QSTVTKNTTTTQTPSKPTTKQKQKPPNPNNDPHEFVFNVPVCSICSNNPCTCAIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 179  
 DB 127 KRIPNKKPGKTTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 246  
 QY 180 TLLTNNNTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSP 226  
 DB 247 FLTSNTARNPELTSTQMETFSTSEGNLSPSQVSTTSEHPSQSPSPNTR 297  
 Query Match 41.1%; Score 505.5; DB 1; Length 292;  
 Best Local Similarity 47.6%; Pred. No. 3.5e-27;  
 Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;  
 QY 1 HKVLTLLAIQDANSQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDANSQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKSNL 126

QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPNPNNDPHEFVFNVPVCSICSNNPCTCAIC 120  
 DB 127 HHTTAQTKGRITTTQTNKPKSKSKNPPKPKDDYHFEVFNVPVCSICGNQLCKSIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 179  
 DB 187 KTIPSNKKPKKPTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 246  
 QY 180 TLLTNNNTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPSQSPSP 226  
 DB 247 SQSTVLDITTPKYTIQQOQLSHSTTSEPTSSQIPTASE-PS-TSPN 291  
 RESULT 11  
 MGNZ18  
 major surface glycoprotein G - human respiratory syncytial virus (strain 18537)  
 C/Species: human respiratory syncytial virus  
 C/Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C/Accession: B32703  
 R/Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987  
 A/Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and B:  
 A/Reference number: A32703; MUID:87289657; PMID:2441388  
 A/Accession: B32703  
 A/Molecule type: mRNA  
 A/Residues: 1-292 <JOH>  
 A/Cross-references: GB:M17213; NID:G333942; PIDN:AAA47412.1; PID:G333943  
 C/Superfamily: respiratory syncytial virus major surface glycoprotein G  
 C/Keywords: glycoprotein; transmembrane protein  
 F:41-63/Domain: transmembrane #status predicted <TMN>  
 F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.4%; Score 497; DB 1; Length 292;  
 Best Local Similarity 46.8%; Pred. No. 1.3e-26;  
 Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;  
 QY 1 HKVLTLLAIQDANSQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVLTLLAIQDANSQIKNTTPTLTQDPQLGIFSNLSITSTQTTILASTTPGVKSNL 126  
 QY 61 QPTVTKNTTTTQTPSKPTTKQKQKPPNPNNDPHEFVFNVPVCSICSNNPCTCAIC 120  
 DB 127 HHTTAQTKGRITTTQTNKPKSKSKNPPKPKDDYHFEVFNVPVCSICGNQLCKSIC 186  
 QY 121 KRIPNKKPGKTTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 179  
 DB 187 KTIPSNKKPKKPTTKPTKPTTKTT-KKDLKPQTTKPKVEPTTKTEPTINTTKNITT 246  
 QY 180 TLLTNNNTGNPKLTSQMETFSTSEGNLSPSQVSTTSEHPS 221  
 DB 247 SQSTVLDITTPKYTIQQOQLSHSTTSEPTSSQIPTASE-PS 287  
 RESULT 12  
 VHNZ  
 nucleocapsid protein (version 2) - human respiratory syncytial virus  
 C/Species: human respiratory syncytial virus  
 C/Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 16-Jul-1999  
 C/Accession: A04026  
 R/Elango, N.; Venkatesan, S.  
 Nucleic Acids Res. 11, 5941-5951, 1983  
 A/Title: Amino acid sequence of respiratory syncytial virus capsid protein.  
 A/Reference number: A04026; MUID:83299261; PMID:6310521  
 A/Accession: A04026  
 A/Molecule type: mRNA  
 A/Residues: 1-467 <ELA>  
 A/Cross-references: GB:X00001; NID:G61215; PIDN:CAA24906.1; PID:G61216  
 C/Genetics: N  
 A/Gene: N  
 C/Superfamily: respiratory syncytial virus nucleocapsid protein  
 C/Keywords: nucleocapsid  
 Query Match 34.6%; Score 425; DB 1; Length 467;



Search completed: October 29, 2003, 17:41:50  
Job time : 23.3245 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:41:00 ; Search time 40.7094 Seconds  
(without alignments)  
975.349 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229  
Sequence: 1 HKVLTATIIQDASQIKNT.....VSTTSEHPQSPSPNTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 642050 seqs, 171146064 residues

Total number of hits satisfying chosen parameters: 642050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	15.7	5179	9	US-09-922-217-1068
2	193	15.7	5179	10	US-09-833-263-1068
3	193	15.7	5179	14	US-10-025-380-1068
4	169.5	13.8	1367	10	US-09-801-368-108
5	162	13.2	941	14	US-10-124-557-14
6	162	13.2	1022	14	US-10-124-557-84
7	162	13.2	1038	14	US-10-124-557-74
8	162	13.2	1049	14	US-10-124-557-58
9	162	13.2	1140	14	US-10-124-557-104
10	162	13.2	1270	14	US-10-124-557-42
11	162	13.2	1311	14	US-10-124-557-42
12	162	13.2	1313	14	US-10-124-557-142
13	162	13.2	1314	14	US-10-124-557-50
14	162	13.2	1320	14	US-10-124-557-46
15	162	13.2	1320	14	US-10-124-557-60

16	162	13.2	1354	14	US-10-124-557-48	Sequence 48, Appl
17	162	13.2	1361	14	US-10-124-557-40	Sequence 40, Appl
18	162	13.2	1363	14	US-10-124-557-52	Sequence 52, Appl
19	162	13.2	1404	14	US-10-124-557-2	Sequence 2, Appl
20	162	13.2	1404	14	US-10-124-557-62	Sequence 62, Appl
21	161.5	13.1	288	9	US-09-216-393-341	Sequence 341, Appl
22	161.5	13.1	288	9	US-09-216-393-344	Sequence 344, Appl
23	161.5	13.1	288	12	US-10-321-856-341	Sequence 341, Appl
24	161.5	13.1	288	12	US-10-321-856-344	Sequence 344, Appl
25	161.5	13.1	800	12	US-10-029-386-32198	Sequence 32198, A
26	152.5	12.4	538	12	US-10-038-694-3	Sequence 3, Appl
27	146	11.9	528	12	US-09-840-746-20	Sequence 20, Appl
28	144.5	11.8	1260	15	US-10-245-802-8	Sequence 8, Appl
29	144.5	11.8	2828	10	US-09-985-129-21	Sequence 21, Appl
30	144.5	11.8	2828	10	US-09-991-630-21	Sequence 21, Appl
31	144.5	11.8	2828	12	US-10-301-822-49	Sequence 49, Appl
32	144.5	11.8	2828	12	US-10-032-189-126	Sequence 126, Appl
33	144.5	11.8	2828	15	US-10-176-847-54	Sequence 54, Appl
34	144.5	11.8	2828	15	US-10-177-293-110	Sequence 110, Appl
35	143	11.6	995	11	US-09-984-130-48	Sequence 48, Appl
36	143	11.6	995	12	US-09-836-353A-48	Sequence 48, Appl
37	142.5	11.6	449	12	US-10-137-870-224	Sequence 224, Appl
38	142.5	11.6	449	12	US-10-140-018-224	Sequence 224, Appl
39	142.5	11.6	449	12	US-10-140-021-224	Sequence 224, Appl
40	142.5	11.6	449	12	US-10-140-274-224	Sequence 224, Appl
41	142.5	11.6	449	12	US-10-140-471-224	Sequence 224, Appl
42	142.5	11.6	449	12	US-10-140-807-224	Sequence 224, Appl
43	142.5	11.6	449	12	US-10-140-922-224	Sequence 224, Appl
44	142.5	11.6	449	12	US-10-140-924-224	Sequence 224, Appl
45	142.5	11.6	449	12	US-10-140-926-224	Sequence 224, Appl

#### ALIGNMENTS

#### RESULT 1

```
US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSTICS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1068
```

Query Match 15.7%; Score 193; DB 9; Length 5179;  
Best Local Similarity 30.8%; Pred. No. 1.6e-06;  
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIQDA--TSQIKNTTP-----TYLQDQDQGLGFSNLSKIT--SQTITLASTP 55  
DB 1504 MTTFIPASTTTLPTTFSPPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563







```

; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-124-557-84

Query Match 13.2%; Score 162; DB 14; Length 1022;
Best Local Similarity 24.6%; Pred. No. 7.8e-05;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIQDASQIKNTTPTVLTQDQIGISFNSLSEITSTTILASTTPGVKSNLQPTTV 65
Db 206 TSKVLAKPTPKAETTK-----GPALTTPKEPTTPPKPEASTTP---KEPTPTTI 253
QY 66 K-----TKNTTTTQTPSKPTTKQKPNKPNNDHFVEVNFVPCISCN 112
Db 254 KSAPTPKPEAPTTTKSAPTTKPEAPTTTKPEAPTTKPEAPTTKPEAPTTKPE 302
QY 113 NPTCWAICKRIPN-----KKFGKKTITKP-----TKKPTFKTK 146
Db 303 APTTTKSAPTTKPEAPTTKPKDAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTK 362
QY 147 KDL-----KPQTTKPEVPTTKP-----TEPTINTKNTITLLTNNITGNPKLTS 194
Db 363 EPAPTAKPKAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKKSAPTTKPEAPT 422
QY 195 QMETFHTSSEGNLSFQSVSTTSEHPSQSPSPNTT 230
Db 423 TTKSAPTTKE-----PSPTTKPEAPTTKPEAPTT 454

RESULT 7
US-10-124-557-74
; Sequence 74, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991

```

```

; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1038 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match 13.2%; Score 162; DB 14; Length 1038;
Best Local Similarity 24.6%; Pred. No. 7.9e-05;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIQDASQIKNTTPTVLTQDQIGISFNSLSEITSTTILASTTPGVKSNLQPTTV 65
Db 190 TSKVLAKPTPKAETTK-----GPALTTPKEPTTPPKPEASTTP---KEPTPTTI 237
QY 66 K-----TKNTTTTQTPSKPTTKQKPNKPNNDHFVEVNFVPCISCN 112
Db 238 KSAPTTKPEAPTTTKSAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKPE 286
QY 113 NPTCWAICKRIPN-----KKFGKKTITKP-----TKKPTFKTK 146
Db 287 APTTTKSAPTTKPEAPTTKPKDAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTK 346
QY 147 KDL-----KPQTTKPEVPTTKP-----TEPTINTKNTITLLTNNITGNPKLTS 194
Db 347 EPAPTAKPKAPTTKPEAPTTKPEAPTTKPEAPTTKPEAPTTKKSAPTTKPEAPT 406
QY 195 QMETFHTSSEGNLSFQSVSTTSEHPSQSPSPNTT 230
Db 407 TTKSAPTTKE-----PSPTTKPEAPTTKPEAPTT 438

RESULT 8
US-10-124-557-58
; Sequence 58, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557

```

```

1 FILING DATE: 16-APR-2002
2 CLASSIFICATION: <UNKNOWN>
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/643,502
5 FILING DATE: 18-JAN-1991
6 APPLICATION NUMBER: US 07/546,114
7 FILING DATE: 29-JUN-1990
8 APPLICATION NUMBER: US 07/457,196
9 FILING DATE: 29-DEC-1989
10 APPLICATION NUMBER: US 07/390,901
11 FILING DATE: 08-AUG-1989
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Cseri, Luann
14 REGISTRATION NUMBER: 31,822
15 REFERENCE/DOCKET NUMBER: GI 5190
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (617) 876-1170
18 TELEFAX: (617) 876-5851
19 INFORMATION FOR SEQ ID NO: 58:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 1049 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 SEQUENCE DESCRIPTION: SEQ ID NO: 58:
26 US-10-124-557-58

```

```

RESULT 9
US-10-124-557-104
Sequence 104, Application US/10124557
Publication No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacob, Kenneth
Hewner, Thomas G.
Geener, Rodney
TITLE OF INVENTION: Megakaryocy
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institi
STREET: 87 CambridgePark I
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 104:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-10-124-557-104

Query Match      13.2%; Score 162; DB 14; Length
Best Local Similarity 24.6%; Pred. No. 8.9e-05;
Matches 68; Conservative 26; Mismatches 104; Indels

QY      6  TTAIIQDATSOIKNTTPTVLQDQLGSPSNLSGILTSQTTTILAS
Db      324  TSKVLAKETPKAETTK-----GPAITPKPTPTTKPEPAS
QY      66  K-----TKNTTTTQTQSPKPTTKQKQKPNKPNNDFH
Db      372  KSAPTTPKEPAPTTTKSAPTTTKPEAPTTKEPAPTTKEPAPTTKEP
QY      113  NPTCWAICKRIPN-----KPKGKTTTKP
Db      421  APTTTKSAPTTTKPEAPTTTPKPAPTTPKPEAPTTTPKEPTPTTKPE
QY      147  KDL-----KPTQTKKPVPTTKP-----TEEPTINTKINITIT
Db      481  EAPATPKKPAPTTPKEAPTTTPKEAPTTTPKEAPTTTPKEAPTTTP
QY      195  QMETFTSHSSSEGNLSPSQVSTTSEHPGQSPSPNWT 230
Db      541  TTKSAPTTTPKE-----PSPTTKPEAPTTTPKEAPTT 572

```

```

RESULT 10
US-10-124-557-44
; Sequence 44, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
;
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
;

```

STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/124,557  
 FILING DATE: 16-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/643,502  
 FILING DATE: 18-JAN-1991  
 APPLICATION NUMBER: US 07/546,114  
 FILING DATE: 29-JUN-1990  
 APPLICATION NUMBER: US 07/457,196  
 FILING DATE: 29-DEC-1989  
 APPLICATION NUMBER: US 07/390,901  
 FILING DATE: 08-AUG-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cseri, Luann  
 REGISTRATION NUMBER: 31,822  
 REFERENCE/DOCKET NUMBER: GI 5190  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)876-1170  
 TELEFAX: (617)876-5851  
 INFORMATION FOR SEQ ID NO: 44:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1270 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 44:

Query Match 13.2%; Score 162; DB 14; Length 1270;  
 Best Local Similarity 24.6%; Pred. No. 0.0001;  
 Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 6 TTAIIQDATSQIKNTTPTTLTQDQGLGSPSNLSEITSTTILASTTPGVKSNLQPTTV 65  
 Db 190 TSKVLAKPTKPAETTK-----GPAULTTKEPTTTPKEPASTTP---KEPTPTTI 237  
 QY 66 K-----TKNTTTQTOPSKPTTKQKQKPPKPNNDFFHEVFNFVPCISCN 112  
 Db 238 KSAPTTPKSPAPTTTKSAPTTPKPAETTTTKEPAPTTPKPEP-----APTTPKEP 286  
 QY 113 NPTCWAICKRIEN-----KKPKKTTTKP-----TKKPTTKTK 146  
 Db 287 APTTKSAPTTPKPAETTTKKAAPTTPKPAETTPKEPTTPKPAETTPKPAETTPK 346  
 QY 147 KDL-----KPQTTKKEVPTTKP-----TEPTINTTKNTITLLTNTTGNPKLTS 194  
 Db 347 EPAPTAPKKAAPTTPKPAETTPKPAETTTKEPSPPTTPKPAETTTKSAPTTPKPAET 406  
 QY 195 QMETFHSTSSSEGNLSQSVSTTSEHPSPSPNNTT 230  
 Db 407 TTKSAPTTPKE-----PSPTTKKPAETTPKPAETTT 438

RESULT 11

US-10-124-557-42  
 ; Sequence 42, Application US/10124557  
 ; Publication No. US20020137894A1  
 ; GENERAL INFORMATION:

APPLICANT: Turner, Katherine  
 Clark, Stephen C.  
 Jacobs, Kenneth  
 Hewick, Rodney M.  
 Gesner, Thomas G.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genetics Institute, Inc.  
 STREET: 87 CambridgePark Drive  
 CITY: Cambridge  
 STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/124,557  
 FILING DATE: 16-Apr-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/643,502  
 FILING DATE: 18-JAN-1991  
 APPLICATION NUMBER: US 07/546,114  
 FILING DATE: 29-JUN-1990  
 APPLICATION NUMBER: US 07/457,196  
 FILING DATE: 29-DEC-1989  
 APPLICATION NUMBER: US 07/390,901  
 FILING DATE: 08-AUG-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cseri, Luann  
 REGISTRATION NUMBER: 31,822  
 REFERENCE/DOCKET NUMBER: GI 5190  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)876-1170  
 TELEFAX: (617)876-5851  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1311 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 42:

US-10-124-557-42

Query Match 13.2%; Score 162; DB 14; Length 1311;  
 Best Local Similarity 24.6%; Pred. No. 0.0001;  
 Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;  
 QY 6 TTAIIQDATSQIKNTTPTTLTQDQGLGSPSNLSEITSTTILASTTPGVKSNLQPTTV 65  
 Db 231 TSKVLAKPTKPAETTK-----GPAULTTKEPTTTPKEPASTTP---KEPTPTTI 278  
 QY 66 K-----TKNTTTQTOPSKPTTKQKQKPPKPNNDFFHEVFNFVPCISCN 112  
 Db 279 KSAPTTPKSPAPTTTKSAPTTPKPAETTTTKEPAPTTPKPEP-----APTTPKEP 327  
 QY 113 NPTCWAICKRIEN-----KKPKKTTTKP-----TKKPTTKTK 146  
 Db 328 APTTKSAPTTPKPAETTTKKAAPTTPKPAETTPKEPTTPKPAETTPKPAETTPK 387  
 QY 147 KDL-----KPQTTKKEVPTTKP-----TEPTINTTKNTITLLTNTTGNPKLTS 194  
 Db 388 EPAPTAPKKAAPTTPKPAETTPKPAETTTKEPSPPTTPKPAETTTKSAPTTPKPAET 447  
 QY 195 QMETFHSTSSSEGNLSQSVSTTSEHPSPSPNNTT 230  
 Db 448 TTKSAPTTPKE-----PSPTTKKPAETTPKPAETTT 479

RESULT 12

US-10-124-557-142  
 ; Sequence 142, Application US/10124557  
 ; Publication No. US20020137894A1  
 ; GENERAL INFORMATION:

APPLICANT: Turner, Katherine





```
QY 113 NPTCWAICKRIPN-----KKPGKXITTKP-----TKKPTFKTK 146
Db ||| ||| ||| : ||| : |||
QY 337 APTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPK 396
Db : ||| ||| ||| : ||| : |||
QY 147 KDL-----KPQTYXPKVEPTTKP-----TEPTINTKTNTITLLTNNITGNPKLTS 194
Db : ||| ||| ||| : ||| : |||
QY 195 QMETFHSTSGNLSPSQVSTTSEHPSOPSPPNNTT 230
Db : : ||| : ||| : |||
QY 457 TTKSAPTTPKE-----PSPTTIKEPAPTTPKEPAPT 488
```

Search completed: October 29, 2003, 17:54:27  
Job time : 41.7094 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:36:05 ; Search time 12.6943 Seconds  
(without alignments)  
773.267 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTAIQDATSIKNT.....VSTTSHPSQSPSPNTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1229	100.0	298	2	US-08-467-963C-8
2	1229	100.0	298	2	US-08-838-189D-8
3	1229	100.0	298	3	US-08-852-344D-8
4	1229	100.0	298	3	US-08-344-639E-8
5	1229	100.0	298	3	US-08-467-969A-8
6	1229	100.0	298	3	US-08-467-961A-8
7	1229	100.0	298	3	US-08-001-554A-8
8	901	73.3	681	6	5194595-19
9	222	18.1	37	3	US-08-793-792-12
10	193	15.7	32	3	US-08-793-792-8
11	188	15.3	216	3	US-08-928-361B-8
12	188	15.3	216	4	US-08-588-995A-8
13	188	15.3	1837	4	US-08-928-361B-5
14	188	15.3	1837	4	US-08-588-995A-5
15	187.5	15.3	1721	3	US-08-700-651-5
16	187.5	15.3	1721	3	US-08-928-361B-6
17	187.5	15.3	1721	4	US-08-588-995A-6
18	181.5	14.8	216	3	US-08-928-361B-27
19	174	14.2	28	3	US-08-793-792-4
20	173.5	14.1	249	3	US-08-700-651-15
21	173.5	14.1	249	3	US-08-928-361B-20
22	173.5	14.1	249	4	US-08-588-995A-20
23	170.5	13.9	263	5	PCT-US91-08177-13
24	164.5	13.4	2476	2	US-08-276-967-2
25	162	13.2	941	4	US-07-757-022B-14
26	162	13.2	1022	4	US-07-757-022B-84
27	162	13.2	1038	4	US-07-757-022B-74

Sequence 58, Appl  
Sequence 104, Appl  
Sequence 44, Appl  
Sequence 42, Appl  
Sequence 142, Appl  
Sequence 50, Appl  
Sequence 46, Appl  
Sequence 60, Appl  
Sequence 48, Appl  
Sequence 40, Appl  
Sequence 52, Appl  
Sequence 2, Appl  
Sequence 62, Appl  
Sequence 341, Appl  
Sequence 344, Appl  
Sequence 12, Appl  
Sequence 17, Appl  
Sequence 17, Appl

## ALIGNMENTS

### RESULT 1

US-08-467-963C-8

; Sequence 8, Application US/08467963C

; Patent No. 5968776

; GENERAL INFORMATION:

; APPLICANT: KLEIN, Michel H

; APPLICANT: DU, Run-Pan

; APPLICANT: EWASYSHYN, Mary E

; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6th Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,963C

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/838,189

; FILING DATE: 16-APR-1997

; APPLICATION NUMBER: US 08/001,554

; FILING DATE: 06-JAN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9200117.1

; FILING DATE: 06-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 298 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-467-963C-8

Query Match	100.0%	Score 1229;	DB 2;	Length 298;
Best Local Similarity	100.0%;	Prod. No. 3e-105;		
Matches 232; Conservative	0;	Mismatches	0;	Indels
Gaps	0;			
1 HKVLTITAIIOQATSIQIKNTTPTYLTDQPQLGISFNSLSEITSQTTLASTTTGVSXNL	60			
67 HKVLTITAIIOQATSIQIKNTTPTYLTDQPQLGISFNSLSEITSQTTLASTTTGVSXNL	126			
61 QPVTVKTKNTTTTTQTQSPKPTTKQRQNPKPNNDHFEVFNPVPCISCSNNPCTCAIC	120			
127 QPVTVKTKNTTTTTQTQSPKPTTKQRQNPKPNNDHFEVFNPVPCISCSNNPCTCAIC	186			
121 KRIENKKPGKTIKTPTKKPFKTKTKDLKPQTKPKKEVPPTKTEBPTNTTKNTIITTT	180			
187 KRIENKKFGKTIKTPIKKPFKTKTKDLKPQTKPKKEVPPTKTEBPTNTTKNTIITTT	246			
181 LLTNNTTGNNPKLTSGMETFHSTSEGNLSPSQVSTTSHEHPSQSPSPNTTRQ	232			
247 LLTNNTTGNNPKLTSGMETFHSTSEGNLSPSQVSTTSHEHPSQSPSPNTTRQ	298			

```

RESULT 2
US-08-838-189D-8
; Sequence 8, Application US/08838189D
; Patent No. 5998169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSHYN, Mary E
; TITLE OF INVENTION: CHIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: MULTIMERIC PROTEIN WHICH CONFERES PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jfb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-838-189D-8

```

Query Match	100.0%;	Score 1229;	DB 2;	Length 298;
Best Local Similarity	100.0%;	Pred. No. 3e-105;		

	Matches	232;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	HKVTLTAAIQDATSQIKNTTPTVLTQDPOLGISFSNLSEITSTWILASTTPGVKSNL	60							
Db	67	HKVTLTAAIQDATSQIKNTTPTVLTQDPOLGISFSNLSEITSTWILASTTPGVKSNL	126							
Qy	61	QPTTVKTKNTTTTQTSKPTTKQKQKPNKPNNDHFEVFNVPVCSIGSNNPTCWAIC	120							
Db	127	QPTTVKTKNTTTTQTSKPTTKQKQKPNKPNNDHFEVFNVPVCSIGSNNPTCWAIC	186							
Qy	121	KRIPNKKPGKNTTKPTKXPTFKTTIKDLKQPTTKPKVEVPTTKTEPTTINTTKNITTT	180							
Db	187	KRIPNKKPGKNTTKPTKXPTFKTTIKDLKQPTTKPKVEVPTTKTEPTTINTTKNITTT	246							
Qy	181	LLTNNTTGNPKLTSOMETFTSTSEGNLSPQSVYTTSEHPQSQSSPPNTTTRQ	232							
Db	247	LLTNNTTGNPKLTSOMETFTSTSEGNLSPQSVYTTSEHPQSQSSPPNTTTRQ	298							

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;

QY 1 HKVLT"TAIQDATSIKNTTPVILTQDPQLGISFNSLSEITSQTTTILASTTGVKSNL 60

Db 67 HKVLTATTAIQQDQATSOIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 126  
QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNPTCWAIC 120  
Db 127 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKKTTPKTKPTKPTTKKDLKPQTKPKKEVPTTKTEBPTNTTKNTTTT 180  
Db 187 KRIPNKKPGKKTTPKTKPTKPTTKKDLKPQTKPKKEVPTTKTEBPTNTTKNTTTT 246  
QY 181 LLTNNNTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
Db 247 LLTNNNTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 4

US-08-344-639E-8  
; Sequence 8, Application US/08344639E  
; Patent No. 6033668  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS  
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS  
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,639E  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-391 MTS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-344-639E-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTATTAIQQDQATSOIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 60  
Db 67 HKVLTATTAIQQDQATSOIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 126  
QY 61 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNPTCWAIC 120  
Db 127 QPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDFFHFEVFNVPVCSICSNPTCWAIC 186  
QY 121 KRIPNKKPGKKTTPKTKPTKPTTKKDLKPQTKPKKEVPTTKTEBPTNTTKNTTTT 180  
Db 187 KRIPNKKPGKKTTPKTKPTKPTTKKDLKPQTKPKKEVPTTKTEBPTNTTKNTTTT 246  
QY 181 LLTNNNTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
Db 247 LLTNNNTGNPKLTSQMETFHSTSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298

## RESULT 5

US-08-467-969A-8  
; Sequence 8, Application US/08467969A  
; Patent No. 6168786  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyszyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,969A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bh  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-969A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTATTAIQQDQATSOIKNTTPTVLTQDPQLGISFNSLSEITSTQTTILASTTTPGVKSNL 60

Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246  
QY 181 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSPSSPNTTRQ 232  
Db 247 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSPSSPNTTRQ 298

## RESULT 6

US-08-467-961A-8  
; Sequence 8, Application US/08467961A  
; Patent No. 6171783  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyshtyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6TH Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,961A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/001,554  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9200117.1  
; FILING DATE: 06-JAN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-476 MIS:bb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-467-961A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 60  
|||||

Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246  
QY 181 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSPSSPNTTRQ 232  
Db 247 LTNNTTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSPSSPNTTRQ 298

## RESULT 7

US-08-001-554A-8  
; Sequence 8, Application US/08001554A  
; Patent No. 6225031  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Michel H  
; APPLICANT: Du, Run-Pan  
; APPLICANT: Ewasyshtyn, Mary E  
; TITLE OF INVENTION: Chimeric Immunogens  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 330 University Avenue, 6th Floor  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/001,554A  
; FILING DATE: 06-JAN-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-286  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; TELEX: 065-24567 SIMBAS  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
US-08-001-554A-8

Query Match 100.0%; Score 1229; DB 3; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3e-105;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 60  
Db 67 HKVLTALIQDATSQIKNTTPTYLTDQPLGIFSLSNLSITSTQTTILASTTPGVKSNL 126  
QY 61 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 120  
Db 127 OPTTVKTKNTTTTQTPSKPTTKQKQKPNKPNNDHFHFEVFNVPSCISNNPTCWAIC 186  
QY 121 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 180  
Db 187 KRIPNKKPGKTKTKPTTKKPTTKKDLKPQTTKPKKVPPTTKPTTEPTINTTKNITTT 246

QY 181 LLTNTTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 232  
|||||  
Db 247 LLTNTTGNPKLTSQMETFHTSTSEGNLSPSQVSTTSEHPSQSPSPNTTQ 298  
|||||

## RESULT 8

5194595-19  
; Patent No. 5194595  
; APPLICANT: WATHEN, MICHAEL W.  
; TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING  
; IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY  
; SYNCYTIAL VIRUS  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/543,780  
; FILING DATE: 31-OCT-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 137,387  
; FILING DATE: 23-DEC-1987  
; SEQ ID NO:19:  
; LENGTH: 681  
5194595-19

Query Match 73.3%; Score 901; DB 6; Length 681;  
Best Local Similarity 91.0%; Pred. No. 1.2e-74;  
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 30 QLGISPSNLSEITSTQTTILASTTPGVKSLQPTTKNTNTTQTQPSKPTTKQKNKP 89  
|||||  
Db 490 QLGISPSNLSEITSTQTTILASTTPGVKSLQPTTKNTNTTQTQPSKPTTKQKNKP 549  
|||||  
QY 90 PNKPNDFHEVFNVPSCSNNPTCWAICRIPNKKPKTKTKTKTKTKTKTKTKTKTK 149  
|:|||||  
Db 550 PSKPNDFHEVFNVPSCSNNPTCWAICRIPNKKPKTKTKTKTKTKTKTKTKTKTK 609  
|:|||||  
QY 150 KPQTTKEVPTTKPTKEPTINTTKNTTTLTNTTGNPKLTSQMETFHTSTSEGNLS 209  
|||||  
Db 610 KPQTTKEVPTTKPTKEPTINTTKNTTTLTNTTGNPKLTSQMETFHTSTSEGNLS 669  
|:|||||  
QY 210 PSQVSTTSE 218  
|||||  
Db 670 PSQVSTTSE 678  
|||||

## RESULT 9

US-08-793-792-12  
; Sequence 12, Application US/08793792  
; Patent No. 6077511  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; TITLE OF INVENTION: Antigenic peptides derived from the  
; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
; NUMBER OF SEQUENCES: 13  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,792  
; FILING DATE:  
; CLASSIFICATION: 514  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
US-08-793-792-12

Query Match 18.1%; Score 222; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 6.3e-14;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 NKPNDHFHEVFNVPSCSNNPTCWAICRIPNKK 127  
|||||  
Db 1 NKPNDHFHEVFNVPSCSNNPTCWAICRIPNKK 37  
|||||

## RESULT 10

US-08-793-792-8  
; Sequence 8, Application US/08793792  
; Patent No. 6077511  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; TITLE OF INVENTION: Antigenic peptides derived from the  
; TITLE OF INVENTION: G-protein of RSV for type- and subtype-specific diagnosis  
; TITLE OF INVENTION: of respiratory syncytial virus (RSV) infection.  
; NUMBER OF SEQUENCES: 13  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,792  
; FILING DATE:  
; CLASSIFICATION: 514  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
US-08-793-792-8

Query Match 15.7%; Score 193; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.4e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 KPNDFHEVFNVPSCSNNPTCWAICR 123  
|||||  
Db 1 KPNDFHEVFNVPSCSNNPTCWAICR 32  
|||||

## RESULT 11

US-08-928-361B-8  
; Sequence 8, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B

Query Match 15.3%; Score 188; DB 3; Length 216;  
Best Local Similarity 26.4%; Pred. No. 8.3e-10;  
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;  
QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPT 63  
Db 12 TTTTITTT 71  
QY 64 TVTKNTKTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNPCISNNPTCWAICKRI 123  
Db 72 TTTTITTT 111  
QY 124 PNKKPGKKTTKTKPTTKTKKOLKPKQTTKPEKVPPTTKPTBPTINTKNTITLLT 183  
Db 112 TTKKPTTT 171  
QY 184 NNTGNPKLTSMETFTSTSEGNLSPSQVSTTSEHPSPSSPNTT 230  
Db 172 TTTTITTTTTTATTTTKKPTTT-----TTTTTTTTKKPTTTTATTTT 214

US-08-928-361B-8

Query Match 15.3%; Score 188; DB 3; Length 216;  
Best Local Similarity 26.4%; Pred. No. 8.3e-10;  
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;  
QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPT 63  
Db 12 TTTTITTT 71  
QY 64 TVTKNTKTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNPCISNNPTCWAICKRI 123  
Db 72 TTTTITTT 111  
QY 124 PNKKPGKKTTKTKPTTKTKKOLKPKQTTKPEKVPPTTKPTBPTINTKNTITLLT 183  
Db 112 TTKKPTTT 171  
QY 184 NNTGNPKLTSMETFTSTSEGNLSPSQVSTTSEHPSPSSPNTT 230  
Db 172 TTTTITTTTTTATTTTKKPTTT-----TTTTTTTTKKPTTTTATTTT 214

RESULT 12  
US-09-588-995A-8  
Sequence 8, Application US/09588995A  
Patent No. 6514697  
GENERAL INFORMATION:  
APPLICANT: PETERSEN, CAROLYN  
APPLICANT: BARNES, DEBRA A.  
APPLICANT: NELSON, RICHARD C.  
APPLICANT: GUT, JIRI  
TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND  
TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM  
TITLE OF INVENTION: INFECTIONS  
FILE REFERENCE: 480.19-5  
CURRENT APPLICATION NUMBER: US/09/588,995A  
CURRENT FILING DATE: 2000-06-06  
PRIORITY APPLICATION NUMBER: 08/827,171  
PRIORITY FILING DATE: 1997-03-27  
PRIORITY APPLICATION NUMBER: 08/928,361  
PRIORITY FILING DATE: 1997-09-12  
PRIORITY APPLICATION NUMBER: 08/700,651  
PRIORITY FILING DATE: 1996-08-14  
PRIORITY APPLICATION NUMBER: 08/415,751  
PRIORITY FILING DATE: 1995-04-03  
NUMBER OF SEQ ID NOS: 115  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 216  
TYPE: PRT  
ORGANISM: Cryptosporidium parvum  
US-09-588-995A-8

Query Match 15.3%; Score 188; DB 4; Length 216;  
Best Local Similarity 26.4%; Pred. No. 8.3e-10;  
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;  
QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPT 63  
Db 12 TTTTITTT 71  
QY 64 TVTKNTKTTTQTPSKPTTKQKQKPNKPNNDPHEFVFNPCISNNPTCWAICKRI 123  
Db 72 TTTTITTT 111  
QY 124 PNKKPGKKTTKTKPTTKTKKOLKPKQTTKPEKVPPTTKPTBPTINTKNTITLLT 183  
Db 112 TTKKPTTT 171  
QY 184 NNTGNPKLTSMETFTSTSEGNLSPSQVSTTSEHPSPSSPNTT 230  
Db 172 TTTTITTTTTTATTTTKKPTTT-----TTTTTTTTKKPTTTTATTTT 214

US-08-928-361B-5

Sequence 5, Application US/08928361B  
Patent No. 6071518  
GENERAL INFORMATION:  
APPLICANT: Petersen, Carolyn  
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
TITLE OF INVENTION: SPECIES INFECTIONS  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
STREET: 385 Sherman Avenue, Suite 6  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-1840  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,361B  
FILING DATE: 12-SEP-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,062  
FILING DATE: 13-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Verny, Hana  
REGISTRATION NUMBER: 30,518  
REFERENCE/DOCKET NUMBER: 480.76-1 (HV)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-1677  
TELEFAX: 650-324-1677  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1837 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-361B-5

Query Match 15.3%; Score 188; DB 3; Length 1837;  
Best Local Similarity 26.4%; Pred. No. 1.4e-08;  
Matches 60; Conservative 23; Mismatches 120; Indels 24; Gaps 2;  
QY 4 TLTAIIQDATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNLQPT 63  
Db 12 TTTTITTT 71



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2003, 17:11:05 ; Search time 52.966 Seconds  
(without alignments)  
695.249 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTAIQDATSQIKNT.....VSTTSEHPQSPSPPTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1229	100.0	232	AAW96314	Secreted G protein
2	1229	100.0	298	AAW96314	Respiratory syncyt
3	1229	100.0	298	AAW96313	Membrane bound G p
4	1120	91.1	298	AAW96313	Sequence of human
5	1120	91.1	298	AAW96313	HSRV glycoprotein
6	1120	91.1	298	AAW96313	HSRV glycoprotein
7	1120	91.1	298	AAW96313	Respiratory syncyt
8	901	73.3	681	AAW96313	Chimeric human res
9	580	47.2	361	ABP97862	RSV G protein pept

10	572.5	46.6	452	22	AAW96314	Amino acid sequenc
11	564.5	45.9	291	24	ABP97861	RSV G protein pept
12	564.5	45.9	548	24	ABP97863	RSV G protein pept
13	562	45.7	349	17	AAW95660	Streptococcal prot
14	562	45.7	349	22	AAW95660	Amino acid sequenc
15	558	45.4	101	15	AAW98253	RSV subgroup A wil
16	558	45.4	101	17	AAW95610	RSV subgp. A prote
17	558	45.4	101	17	AAW95610	Respiratory Syncyt
18	558	45.4	101	17	AAW95610	Respiratory Syncyt
19	558	45.4	101	20	AAW97050	RSV G protein anti
20	558	45.4	101	21	AAW97050	A G2Na peptide der
21	558	45.4	101	22	AAW97050	Amino acid sequenc
22	558	45.4	101	22	AAW97050	Amino acid sequenc
23	558	45.4	101	22	AAW97050	Amino acid sequenc
24	558	45.4	101	23	AAW97050	Human G protein, G
25	552.5	45.0	356	17	AAW95661	Immunogenic carrie
26	538	43.8	101	16	AAW98255	RSV subgroup A mod
27	538	43.8	101	17	AAW95612	RSV subgp. A prote
28	538	43.8	101	17	AAW95618	Respiratory Syncyt
29	538	43.8	101	17	AAW97052	RSV G protein anti
30	538	43.8	101	20	AAW97052	Peptide which indu
31	538	43.8	101	22	AAW97311	Amino acid sequenc
32	538	43.8	101	22	AAW97311	Human G protein, G
33	538	43.8	101	23	AAW97311	Peptide which indu
34	538	43.8	101	23	AAW97311	Human G protein, G
35	538	43.8	101	23	AAW97311	Human G protein, G
36	534	43.4	101	20	AAW97310	Peptide which indu
37	534	43.4	101	23	AAW97312	Human G protein, G
38	514	41.8	101	20	AAW97312	Peptide which indu
39	514	41.8	101	23	AAW97312	Human G protein, G
40	507	41.3	299	22	AAW96314	Amino acid sequenc
41	506	41.2	101	17	AAW95614	RSV subgp. A prote
42	506	41.2	101	17	AAW97063	Respiratory Syncyt
43	506	41.2	101	20	AAW94090	RSV G protein anti
44	506	41.2	101	22	AAW94090	Amino acid sequenc
45	506	41.2	101	23	AAW94090	Human G protein, G

#### ALIGNMENTS

RESULT 1  
AAW96314  
ID AAW96314 standard; Protein; 232 AA.  
AC AAW96314;  
XX  
XX  
DT 28-JUN-1999 (first entry)  
DE Secreted G protein of respiratory syncytial virus.  
DE G protein; respiratory syncytial virus; RSV; recombinant vector;  
KW vaccine; immune response; immunogenicity; tPA; antibody;  
KW tissue plasminogen activator.  
XX  
XX  
OS Respiratory syncytial virus.  
OS  
XX  
PN WO9904010-A1.  
XX  
XX  
PD 28-JAN-1999.  
XX  
PF 16-JUL-1998; 98WO-CA00697.  
XX  
PR 18-JUL-1997; 97US-0896442.  
XX  
XX (CONN-) CONNAUGHT LAB LTD.  
XX  
XX Klein ME, Li X, Sambhara S;  
XX WPI; 1999-132254/11.  
XX N-PSDB; AAW08422.  
XX  
XX Immunogenic composition for generating antibodies against



PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and  
 PT to raise antibodies for diagnosis  
 XX  
 PS Claim 9; Fig 3; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods. This truncated G  
 CC protein is secreted since it lacks a transmembrane domain.  
 XX  
 SQ Sequence 232 AA;  
 Query Match 100.0%; Score 1229; DB 20; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVTLTALIQDATSQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 DB 1 HKVTLTALIQDATSQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 QY 61 OPTTVTKNTTTTQTPSKPTTKQKPKPKPNKPNDFHEVEFVPCSCSNNTPTCWAIC 120  
 DB 61 OPTTVTKNTTTTQTPSKPTTKQKPKPKPNKPNDFHEVEFVPCSCSNNTPTCWAIC 120  
 QY 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKPKKEVPTTKPTKEPTINTTKNITTT 180  
 DB 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKPKKEVPTTKPTKEPTINTTKNITTT 180  
 QY 181 LNTNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 DB 181 LNTNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 RESULT 2  
 AAR39286  
 ID AAR39286 standard; Protein; 298 AA.  
 XX  
 AC AAR39286;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 13-JAN-1994 (first entry)  
 XX  
 DE Respiratory syncytial virus (RSV) G protein.  
 XX  
 KW PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO9314207-A1.  
 XX  
 PD 22-JUL-1993.  
 XX  
 PF 05-JAN-1993; 93WO-CA00001.  
 XX  
 PR 06-JAN-1992; 92GB-0000117.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Ewasysyn ME, Klein MH;  
 XX  
 DR WPI; 1993-243222/30.  
 DR N-PSDB; AAQ45686.

XX Multimeric hybrid genes and their chimeric proteins - are  
 PT vaccines against multiple pathogenic infections e.g.  
 PT para-influenza virus and respiratory syncytial virus  
 XX  
 PS Claim 11; Figure 7A-7D; 80pp; English.  
 XX  
 CC A novel multimeric hybrid gene is used as a vaccine. The gene  
 CC consists of two gene sequences which are linked and encode antigenic  
 CC regions, these two sequences being derived from two different  
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus  
 CC (RSV)). The gene sequences that are particularly used are those  
 CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and  
 CC G proteins (AAQ45685, AAQ45686).  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 298 AA;  
 Query Match 100.0%; Score 1229; DB 14; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVTLTALIQDATSQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVTLTALIQDATSQIKNTTPTLYLTQDPQLGISFNSLSEITSTQTTILASTTPGVKSNL 126  
 QY 61 OPTTVTKNTTTTQTPSKPTTKQKPKPKPNKPNDFHEVEFVPCSCSNNTPTCWAIC 120  
 DB 127 OPTTVTKNTTTTQTPSKPTTKQKPKPKPNKPNDFHEVEFVPCSCSNNTPTCWAIC 186  
 QY 121 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKPKKEVPTTKPTKEPTINTTKNITTT 180  
 DB 187 KRIPNKKPKGKTTTKTKPTFKTKKOLKPKQTKPKKEVPTTKPTKEPTINTTKNITTT 246  
 QY 181 LNTNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 232  
 DB 247 LNTNNTGNPKLTSQMETFHTSSSEGNLSPSQVSTTSEHPSQSPSPNTTRQ 298  
 RESULT 3  
 AAW96313  
 ID AAW96313 standard; Protein; 298 AA.  
 XX  
 AC AAW96313;  
 XX  
 DT 28-JUN-1999 (first entry)  
 XX  
 DE Membrane bound G protein of respiratory syncytial virus.  
 XX  
 KW G protein; respiratory syncytial virus; RSV; recombinant vector;  
 KW vaccine; immune response; immunogenicity; tPA; antibody;  
 KW tissue plasminogen activator.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 PN WO9904010-A1.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 16-JUL-1998; 98WO-CA00697.  
 XX  
 PR 18-JUL-1997; 97US-0896442.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Klein MH, Li X, Sambhara S;  
 XX  
 DR WPI; 1999-132254/11.  
 DR N-PSDB; AAX08421.  
 XX  
 PT Immunogenic composition for generating antibodies against  
 PT respiratory syncytial virus - comprises non-replicating vector  
 PT containing the protein G sequence, useful in protective vaccines and

PT to raise antibodies for diagnosis  
 XX Claim 4; Fig 2; 67pp; English.  
 XX  
 CC The respiratory syncytial virus (RSV) G protein can be used in  
 CC vaccines by inserting the G protein gene into a non-replicating  
 CC vector. The G protein is placed under the control of alternative  
 CC signal and expression sequences, for example the chimeric G protein  
 CC produced may also comprise the signal peptide of tissue plasminogen  
 CC activator (tPA). The recombinant vector may also comprise sequences  
 CC upstream of the G protein gene which enhance the G proteins  
 CC immunoprotective ability. The resulting immunogenic composition will  
 CC generate antibodies directed against the RSV G protein when  
 CC administered to a host organism. The composition is useful as a  
 CC vaccine to immunise against RSV-associated disease, particularly  
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,  
 CC by usual immunisation and cell fusion methods.  
 XX  
 SQ Sequence 298 AA;  
 Query Match 100.0%; Score 1229; DB 20; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-89;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HKVLTLLIIODATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTGGVKSNL 60  
 DB 67 HKVLTLLIIODATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTGGVKSNL 126  
 QY 61 QPTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 QPTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTKTKPTKTKKDKLKPQTKPKVEVPTTKTEPTINTKNTITTT 180  
 DB 187 KRIPNKKPGKKTTKTKPTKTKKDKLKPQTKPKVEVPTTKTEPTINTKNTITTT 246  
 QY 181 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPNTTRQ 232  
 DB 247 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPNTTRQ 298  
 RESULT 4  
 AAP70845  
 ID AAP70845 standard; protein; 298 AA.  
 XX  
 AC AAP70845;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-APR-1991 (first entry)  
 XX  
 DE Sequence of human respiratory syncytial virus (HRSV) A2 strain  
 DE G protein.  
 XX  
 KW Vaccine.  
 XX  
 XX Human respiratory syncytial virus (HRSV).  
 OS  
 OS WO8704185-A.  
 PN  
 XX  
 PD 16-JUL-1987.  
 XX  
 XX 23-DEC-1986; 86WO-US02756.  
 PF  
 XX 14-JAN-1986; 86US-0818740.  
 PR  
 XX (UYN-) UNIV NORTH CAROLINA.  
 PA (WERTZ) WERTZ G W.  
 XX  
 XX WPI; 1987-206300/29.  
 DR N-PSDB; AAN70784.  
 XX  
 XX Vaccines for human respiratory virus - comprising proteins or  
 PT fragment encoded by a DNA sequence coding for human respiratory

PT syncytial virus proteins.  
 XX Disclosure; Chart 13; 57pp; English.  
 XX  
 CC A novel plasmid which comprises a DNA sequence encoding this  
 CC protein, and the protein itself, are claimed, for use as HRSV  
 CC vaccines. The vaccine can be administered to pregnant women or to  
 CC women of child bearing age to stimulate maternal antibodies.  
 CC Infants can also be vaccinated at 2-3 months of age.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 298 AA;  
 Query Match 91.1%; Score 1120; DB 8; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTLLIIODATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTGGVKSNL 60  
 DB 67 HKVLTLLIIODATSQIKNTTPTLTQDPQLGISFNSLSEITSTQTTILASTTGGVKSNL 126  
 QY 61 QPTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 120  
 DB 127 QPTVTKNTNTTQTQPSKPTTKQKQKPPKPNNDHFEVFNVPVCSICSNPTCWAIC 186  
 QY 121 KRIPNKKPGKKTTKTKPTKTKKDKLKPQTKPKVEVPTTKTEPTINTKNTITTT 180  
 DB 187 KRIPNKKPGKKTTKTKPTKTKKDKLKPQTKPKVEVPTTKTEPTINTKNTITTT 246  
 QY 181 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPNTTRQ 232  
 DB 247 LLTNNITGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSPSPNTTRQ 298  
 RESULT 5  
 AAR25302  
 ID AAR25302 standard; Protein; 298 AA.  
 XX  
 AC AAR25302;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 03-MAR-1993 (first entry)  
 XX  
 DE HRSV glycoprotein G (gpG).  
 XX  
 KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;  
 KW major capsid protein; N.  
 XX  
 OS Human respiratory syncytial virus strain A2.  
 XX  
 PN US5149650-A.  
 XX  
 PD 22-SEP-1992.  
 XX  
 PF 13-JUL-1988; 88US-0218737.  
 XX  
 PR 14-JAN-1986; 86US-0818740.  
 PR 13-JUL-1988; 88US-0218737.  
 XX  
 PA (UYN-) UNIV NORTH CAROLINA.  
 XX  
 XX Collins PL, Wertz GW;  
 PI WPI; 1992-340247/41.  
 DR N-PSDB; AAQ29623.  
 DR  
 XX Vaccines for human respiratory virus - include structural genes  
 PT coding for native structural viral proteins and immunogenic  
 PT fragments  
 XX  
 PS Disclosure; Page 18; 21pp; English.  
 XX  
 CC The sequences of mRNA encoding HRSV structural proteins are given in

CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid  
 CC protein N. The sequences and encoded proteins are useful for  
 CC preparing vaccines against HRSV. The vaccines can be used to confer  
 CC immunity against respiratory tract infections on human subjects.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 298 AA;  
 SQ

Query Match 91.1%; Score 1120; DB 13; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVLTLLAIQDQATSIQKNTTPTLYLTQDPLGIFSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVPTTLLAIQDQATSIQKNTTPTLYLTQDPLGIFSNLSSEITSTQTTILASTTPGVKSNL 126

QY 61 QPTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120  
 DB 127 QSTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186

QY 121 KRIPNKKPGKTTTKPTKPTTKKDLKPQTTKPKVPTTKPTTEPTINTTKNIITT 180  
 DB 187 KRIPNKKPGKTTTKPTKPTTKKDLKPQTTKPKVPTTKPTTEPTINTTKNIITT 246

QY 181 LLTNNNTGNPKLTSMETFHSTSEGNLSPSQSVSTTSEHPSQSPSPPTTRQ 232  
 DB 247 LLTNTTGNPELTSMETFHSTSEGNLSPSQSVSTTSEHPSQSPSPPTTRQ 298

RESULT 6  
 AAU47605  
 ID AAU47605 standard; Protein; 298 AA.  
 XX  
 AC AAU47605;  
 XX  
 DT 11-JUN-1998 (first entry)  
 DE HRSV glycoprotein G.  
 XX  
 KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.  
 XX Human respiratory syncytial virus.  
 OS  
 PN US5716823-A.  
 XX  
 PD 10-FEB-1998.  
 XX  
 PF 12-MAY-1997; 97US-0854783.  
 XX  
 PR 13-JUL-1988; 88US-0218737.  
 PR 14-JAN-1986; 86US-0818740.  
 PR 23-DEC-1986; 86WO-US02756.  
 PR 11-JUN-1992; 92US-0897171.  
 PR 12-MAY-1997; 97US-0854783.  
 XX  
 PA (PRAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Collins PL, Wertz GW;  
 XX  
 DR WPI; 1998-144802/13.  
 DR N-PSDB; AAV18736.  
 XX  
 PT Production of human respiratory syncytial virus glyco-protein F or G  
 PT - by culturing eukaryotic host cells transfected with corresponding  
 PT DNA  
 XX  
 PS Example 1; Columns 27-28; 17pp; English.  
 XX  
 CC The present sequence was used in the development of a novel method  
 CC for the production of human respiratory syncytial virus (HRSV)  
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises  
 CC culturing eukaryotic host cells transfected with an isolated DNA  
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare

CC vaccines against HRSV.  
 XX Sequence 298 AA;  
 SQ

Query Match 91.1%; Score 1120; DB 19; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVLTLLAIQDQATSIQKNTTPTLYLTQDPLGIFSNLSSEITSTQTTILASTTPGVKSNL 60  
 DB 67 HKVPTTLLAIQDQATSIQKNTTPTLYLTQDPLGIFSNLSSEITSTQTTILASTTPGVKSNL 126

QY 61 QPTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 120  
 DB 127 QSTTVTKNTTTTQTPSKPTTKQKONKPPKPNNDHFEVFNFPVCSICSNPTCWAIC 186

QY 121 KRIPNKKPGKTTTKPTKPTTKKDLKPQTTKPKVPTTKPTTEPTINTTKNIITT 180  
 DB 187 KRIPNKKPGKTTTKPTKPTTKKDLKPQTTKPKVPTTKPTTEPTINTTKNIITT 246

QY 181 LLTNNNTGNPKLTSMETFHSTSEGNLSPSQSVSTTSEHPSQSPSPPTTRQ 232  
 DB 247 LLTNTTGNPELTSMETFHSTSEGNLSPSQSVSTTSEHPSQSPSPPTTRQ 298

RESULT 7  
 AAU74676  
 ID AAU74676 standard; Protein; 298 AA.  
 XX  
 AC AAU74676;  
 XX  
 DT 09-APR-2002 (first entry)  
 DE Respiratory syncytial virus G protein.  
 XX  
 KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;  
 KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;  
 KW antiviral chemotherapeutic compound; humoral response;  
 KW cellular immune response; hPIV; paediatric respiratory disease;  
 KW globin gene transfer; sickle cell disease; beta-thalassaemia;  
 KW human immunodeficiency virus infection; HIV.  
 XX Human respiratory syncytial virus.  
 OS  
 PN WO200192548-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 22-MAY-2001; 2001WO-US16610.  
 XX  
 PR 01-JUN-2000; 2000US-208701P.  
 XX  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Portner A, Takimoto T;  
 XX  
 DR WPI; 2002-130534/17.  
 DR N-ESDB; AAS21045.  
 XX  
 PT Recombinant Sendai virus useful in vaccines to protect infection by  
 PT paramyxoviruses, comprises exogenous nucleic acid encoding  
 PT paramyxovirus protein or its antigenic fragment -  
 XX  
 PS Disclosure; Page 46; 57pp; English.  
 XX  
 CC The invention relates to a recombinant Sendai virus comprising an  
 CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its  
 CC antigenic fragment. The virus may be administered in combination  
 CC with an antiviral chemotherapeutic compound. Two or more viruses  
 CC expressing different PMV proteins may be co-administered. Compositions  
 CC comprising the virus are useful for eliciting a humoral and/or  
 CC cellular immune response to a PMV in a mammal, particularly a human.  
 CC Further a recombinant Sendai virus comprising an exogenous nucleic acid

CC encoding a second PMV protein is also administered and priming and/or  
 CC boosting humoral or cellular immune response comprises administering  
 CC one or more of a recombinant or isolated PMV protein or its antigenic  
 CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral  
 CC vector encoding a PMV protein. The recombinant virus is useful as an  
 CC effective vaccine against hPIV or RSV (the major causes of paediatric  
 CC respiratory disease) and also to express any gene of  
 CC interest in target cells, providing a positive medical impact on  
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)  
 CC into stem cells effects a cure for sickle cell disease or beta-  
 CC thalassaemia. The recombinant virus may also prove effective in  
 CC conferring immunity to human immunodeficiency virus (HIV) infection.  
 CC The Sendai virus replicates at level that is high enough to  
 CC induce sufficient immunity, but does not cause any harm to human  
 CC recipient. The present sequence represents a respiratory syncytial  
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein  
 CC suitable for expression by the recombinant virus of the invention.  
 XX Sequence 298 AA;  
 SQ  
 Query Match 91.1%; Score 1120; DB 23; Length 298;  
 Best Local Similarity 92.7%; Pred. No. 1.8e-80;  
 Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 HKVLTITAIIDATSOIKNTPTVLTQDPOLGISFSNLSEITQTITILASTTPGVKSNL 60  
 DB 67 HKVLTITAIIDATSOIKNTPTVLTQDPOLGISFSNLSEITQTITILASTTPGVKSTL 126  
 QY 61 QPTTKKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFPVPCISCSNNPTCWAIC 120  
 DB 127 QSTVTKNTTTTQTPSKPTTKQKQKPNKPNNDPHEFVFPVPCISCSNNPTCWAIC 186  
 QY 121 KRIIPKPKGKTTTKPKTKPTTKKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 180  
 DB 187 KRIIPKPKGKTTTKPKTKPTTKKDLKPQTTKPKKEVPTTKTEPTINTTKNITTT 246  
 QY 181 LLTNTTGNPKLTSOMETHSTSEGNLSPOVSTTSEHPSQSPSPPTTRQ 232  
 DB 247 LLTNTTGNBELTSQMETFHTSSEGNPSPQVSTTSEYFSPSPSPPTPRQ 298  
 RESULT 8  
 AAP90441  
 ID AAP90441 standard; protein; 681 AA.  
 AC AAP90441;  
 DT 01-NOV-1989 (first entry)  
 DE Chimeric human respiratory syncytial virus glycoproteins F and G.  
 DE Chimeric polypeptide; human respiratory syncytial virus;  
 KW protein F; protein G; vaccine.  
 XX Human respiratory syncytial virus.  
 XX WO8905823-A.  
 XX 29-JUN-1989.  
 XX 31-OCT-1988; 88WO-US03784.  
 XX 23-DEC-1987; 87US-0137387.  
 XX (UPJO ) UPJOHN CO.  
 XX Wathen M;  
 XX WPI; 1989-206593/28.  
 XX Chimeric human respiratory syncytial virus polypeptides (s)  
 PT - contg. immunogenic fragments from HRSV glycoproteins  
 PT F and G, for vaccine prodn.

XX Claim 3; page 47-48; 50pp; English.  
 XX Chimeric polypeptide contg. a signal sequence and one or more  
 CC immunogenic fragments from both human respiratory syncytial virus  
 CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg  
 CC E. coli, Chinese hamster ovary cells, murine C127 cells and  
 CC S. frugiperda.  
 XX Sequence 681 AA;  
 SQ  
 Query Match 73.3%; Score 901; DB 10; Length 681;  
 Best Local Similarity 91.0%; Pred. No. 8.6e-63;  
 Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;  
 QY 30 QLGISFSNLSEITQTITILASTTPGVKSNLOPTTKNTTQTOTPSKPTTKQKQKPN 89  
 DB 490 QLGISFSNLSEITQTITILASTTPGVKSTLOSTTKNTTQTOTPSKPTTKQKQKPN 549  
 QY 90 PNKPNNDPHEFVFPVPCISCSNNPTCWAICRIIPKPKGKTTTKPKTKPTTKTKDL 149  
 DB 550 PSKPNNDPHEFVFPVPCISCSNNPTCWAICRIIPKPKGKTTTKPKTKPTTKTKDL 609  
 QY 150 KPQTTKPKKEVPTTKTEPTINTTKNITTTLLTNTTGNPKLTSOMETHSTSEGNLS 209  
 DB 610 KPQTTKPKKEVPTTKTEPTINTTKNITTTLLTNTTGNBELTSQMETFHTSSEGNPS 669  
 QY 210 PSQVSTTSE 218  
 DB 670 PSQVNNISQ 678  
 RESULT 9  
 ABP97862  
 ID ABP97862 standard; protein; 361 AA.  
 AC ABP97862;  
 DT 03-JUN-2003 (first entry)  
 DE RSV G protein peptide G2Na and diphtheria anatoxin peptide DTB.  
 DE Diphtheria anatoxin; immune response; antigen; carrier; G protein;  
 KW hypersensitivity response; vaccine; infection; RSV; cancer.  
 XX Synthetic.  
 OS Respiratory syncytial virus.  
 OS Corynebacterium diphtheriae.  
 XX FR2827606-A1.  
 XX 24-JAN-2003.  
 XX 20-JUL-2001; 2001FR-0009733.  
 XX 20-JUL-2001; 2001FR-0009733.  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX Corvaia N, Nguyen NT, Beck A;  
 XX WPI; 2003-241969/25.  
 XX New peptide derived from diphtheria anatoxin, useful as carrier in  
 PT vaccines, lacks at least one Cys residue, also related nucleic acids -  
 XX Disclosure; Page 27-28; 42pp; French.  
 XX The present sequence represents a fusion protein comprising a peptide  
 CC derived from the G protein of Respiratory syncytial virus (RSV) linked  
 CC to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin  
 CC peptide generates an immune response against any antigen coupled to it,  
 CC and has a negative, immediate hypersensitivity response. Diphtheria

CC anatoxin derived peptides, when modified to lack at least one cysteine  
CC residue, are useful as carrier peptides. Deletion of Cys residues in  
CC anatoxin peptides reduces formation of unwanted disulfide bridges. The  
CC peptides are used as a carrier for vaccines, particularly those for  
CC prevention or treatment of viral, bacterial, parasitic or fungal  
CC infections, or cancers and to generate, or increase, an immune response  
CC against infectious agents or tumour cells.

XX Sequence 361 AA;  
SQ

Query Match 47.2%; Score 580; DB 24; Length 361;  
Best Local Similarity 70.8%; Pred. No. 8.3e-38;  
Matches 114; Conservative 5; Mismatches 20; Indels 22; Gaps 3;

QY 64 TVTKNTNTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCSCSNPTCWAICKRI 123  
Db 4 TVTKNTNTTQTPSKPTTKQKQKPNKPNNDPHEVFVPCSCSNPTCWAICKRI 63  
QY 124 PNKKPGKTTTKTKPTTKTKKDLKPKPTTKPKVPTTKPTEPTIN-----TTKT 175  
Db 64 PNKKPGKTTTKTKPTTKTKKDLKPKPTTKPKVPTTKPTEPTIN-----TTKT 175  
QY 176 NITT-----TLTNNTTGNPKLTSQMETPHSTSE 205  
Db 121 KIESLKEHGPINKMSESPNKTVSEKAKQYLEEFHQTALE 161

RESULT 10  
AAB67771  
ID AAB67771 standard; Protein; 452 AA.  
AC AAB67771;  
XX  
DT 11-JUN-2001 (first entry)  
XX  
DE Amino acid sequence of a fusion protein of P40 and RSV antigen.  
XX  
KW Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;  
KW RSV; RSV infection; lung; respiratory tract; vaccine.  
XX  
OS Synthetic.  
OS Klebsiella pneumoniae.  
XX Respiratory syncytial virus.  
PN WO200121203-A1.  
XX  
PD 29-MAR-2001.  
XX  
PF 22-SEP-2000; 2000WO-FR02626.  
XX  
PR 23-SEP-1999; 99FR-0011888.  
XX  
PA (FABR ) FABRE MEDICAMENT SA PIERRE.  
PI Corvaiea N, Goestch L;  
XX  
DR WPI; 2001-257929/26.  
XX N-PSDB; AAF80153.  
XX  
PT Vaccine against respiratory syncytial virus, comprises enterobacterial  
PT outer membrane protein and viral immunogen, provides protective  
PT response throughout the respiratory tract -  
XX  
PS Example 2; Page 31-32; 39pp; French.  
XX  
CC The present sequence represents a fusion protein comprising a Klebsiella  
CC pneumoniae outer membrane protein A (OmpA) designated P40 and a  
CC respiratory syncytial virus (RSV) antigen. Enterobacterium OmpA proteins,  
CC associated with an immunogenic peptide from RSV are used to prepare  
CC a nasal composition that induces a protective response, against RSV  
CC infection in the upper and lower (lung) respiratory tract. OmpA  
CC potentiates the immune response to some immunogenic peptides,  
CC eliminating the need for adjuvants. The method is useful for producing

CC vaccines for prevention or treatment of RSV infections.

SQ Sequence 452 AA;

Query Match 46.6%; Score 572.5; DB 22; Length 452;  
Best Local Similarity 76.1%; Pred. No. 4.2e-37;  
Matches 108; Conservative 5; Mismatches 18; Indels 11; Gaps 2;

QY 25 LTQDPQLGISFSLSEITTSQITLILASTTPPGVKSNLQPTTVTKTKNTTTTQTPSKPTTKQ 84  
Db 322 LAPDRKRVLEVGKYEVTQ-----PQPG-----DPMVTKTKNTTTTQTPSKPTTKQ 370  
QY 85 RQKPNKPNNDPHEVFVPCSCSNPTCWAICKRIPNKKPGKTTTKPKPTFKT 144  
Db 371 RQKPNKPNNDPHEVFVPCSCSNPTCWAICKRIPNKKPGKTTTKPKPTFKT 430  
QY 145 TKKDLKPKPTTKPKVPTTKPTE 166  
Db 431 TKKDLKPKPTTKPKVPTTKPVD 452

RESULT 11

ABP97861  
ID ABP97861 standard; protein; 291 AA.  
XX

AC ABP97861;

XX  
DT 03-JUN-2003 (first entry)

DE RSV G protein peptide G2Na and diphtheria anatoxin peptide DTa.

XX Diphtheria anatoxin; immune response; antigen; carrier; G protein;  
KW hypersensitivity response; vaccine; infection; RSV; cancer.

XX Synthetic.

OS Respiratory syncytial virus.

OS Corynebacterium diphtheriae.

XX FR2827606-A1.

PN 24-JAN-2003.

PD 20-JUL-2001; 2001FR-0009733.

PF 20-JUL-2001; 2001FR-0009733.

XX (FABR ) FABRE MEDICAMENT SA PIERRE.

PA Corvaiea N, Nguyen NT, Beck A;

PI WPI; 2003-241969/25.

DR New peptide derived from diphtheria anatoxin, useful as carrier in  
XX vaccines, lacks at least one Cys residue, also related nucleic acids -  
XX Disclosure; Page 26-27; 42pp; French.

CC The present sequence represents a fusion protein comprising a peptide  
CC derived from the G protein of Respiratory syncytial virus (RSV) linked  
CC to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin  
CC peptide generates an immune response against any antigen coupled to it,  
CC and has a negative, immediate hypersensitivity response. Diphtheria  
CC anatoxin derived peptides, when modified to lack at least one cysteine  
CC residue, are useful as carrier peptides. Deletion of Cys residues in  
CC anatoxin peptides reduces formation of unwanted disulfide bridges. The  
CC peptides are used as a carrier for vaccines, particularly those for  
CC prevention or treatment of viral, bacterial, parasitic or fungal  
CC infections, or cancers and to generate, or increase, an immune response  
CC against infectious agents or tumour cells.

SQ Sequence 291 AA;

Query Match 45.9%; Score 564.5; DB 24; Length 291;

	Best Local Similarity	67.1%;	Pred. No. 1.1e-36;	
	Matches 110;	Conservative	10; Mismatches 21;	Indels 23; Gaps 3;
QY	64	TVKTKNTTTTQTQPSKPTTKORQNKPNNDFHFEVFNFPVCSICSNNPTCWAICKRI	123	
DBb	4	TVKTKNTTTTQTQPSKPTTKORQNKPNNDFHFEVFNFPVCSICSNNPTCWAICKRI	63	
QY	124	PNKPGKKTTTTPKKTPFTTKTKDKLPQTKPKGVFTTKPTE--EPTINTKTNTITTL	181	
Dbb	64	PNKPGKKTTTTPKKTPFTTKTKDKLPQTKPKGVFTTKPTE:::~::~:	119	
QY	182	LTNNTTGNPKLTSOMETFHSTSEGNLSPSOVSTTSEHPQSQQSS	225	
DBb	120	-----MENF---SSYHGTKGYVDSIQGKIQPKS	146	
RESULT	12			
ID	ABP97863			
AC	ABP97863 standard; protein; 548 AA.			
XX	ABP97863;			
DT	03-JUN-2003 (first entry)			
DE	RSV G protein peptide G2Na and diphtheria anatoxin peptide DtaDb.			
KW	Diphtheria anatoxin; immune response; antigen; carrier; G protein;			
XW	hypersensitivity response; vaccine; infection; RSV; cancer.			
OS	Synthetic.			
OS	Respiratory syncytial virus.			
OS	Corynebacterium diphtheriae.			
EN	PR2827606-A1.			
PD	24-JAN-2003.			
PF	20-JUL-2001; 2001FR-0009733.			
PR	20-JUL-2001; 2001FR-0009733.			
PA	(FABR ) FABRE MEDICAMENT SA PIERRE.			
PI	Corvaia N, Nguyen NT, Beck A;			
DR	WPI; 2003-241969/25.			
PT	New peptide derived from diphtheria anatoxin, useful as carrier in			
PT	vaccines, lacks at least one Cys residue, also related nucleic acids			
XX	Disclosure; Page 29-30; 42pp; French.			
CC	The present sequence represents a fusion protein comprising a peptide			
CC	derived from the G protein of Respiratory syncytial virus (RSV) linked			
CC	to a peptide derived from a diphtheria anatoxin. The diphtheria anatoxin			
CC	peptide generates an immune response against any antigen coupled to it,			
CC	and has a negative, immediate hypersensitivity response. Diphtheria			
CC	anatoxin derived peptides, when modified to lack at least one cysteine			
CC	residue, are useful as carrier peptides. Deletion of Cys residues in			
CC	anatoxin peptides reduces formation of unwanted disulfide bridges. The			
CC	peptides are used as a carrier for vaccines, particularly those for			
CC	prevention or treatment of viral, bacterial, parasitic or fungal			
CC	infections, or cancers and to generate, or increase, an immune response			
CC	against infectious agents or tumour cells.			
XX				
SQ	Sequence 548 AA;			
Query Match	45.9%;	Score 564.5;	DB 24;	Length 548;
Best Local Similarity	67.1%;	Pred. No. 2.2e-36;		
Matches 110;	Conservative	10; Mismatches 21;	Indels	23; Gaps 3;
QY	64	TVKTKNTTTTQTQPSKPTTKORQNKPNNDFHFEVFNFPVCSICSNNPTCWAICKRI	123	

Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

QY 37 NLSEITSQTT--ILASTT-----PGVKSNOPTTVTKTKNTTTTQTPSKPTTKQKQKP 89  
 Db 216 NGKTLGKETTTEAADAATARSFNFPILNSM---TVTKTKNTTTTQTPSKPTTKQKQKP 272  
 QY 90 PNKPNDHFHEVFNFPVPCISCSNNPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 Db 273 PNKPNDHFHEVFNFPVPCISCSNNPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 332  
 QY 150 KPQTTKPKVEPTTKPTE 166  
 Db 333 KPQTTKPKVEPTTKPVD 349

RESULT 14  
 AAB68028  
 ID AAB68028 standard; Protein; 349 AA.  
 XX  
 AC AAB68028;  
 XX  
 DT 29-JUN-2001 (first entry)  
 XX  
 DE Amino acid sequence of fusion protein comprising 2 G protein fragments.  
 XX  
 KW Aliphatic ammonium salt; immunogen; antigen; syncytial virus infection;  
 KW G protein; fusion protein.  
 XX  
 OS Synthetic.  
 OS Streptococcus sp.  
 OS Respiratory syncytial virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 120..230  
 FT /note= "G protein fragment of VRS"  
 XX  
 XX FR2798292-A1.  
 XX  
 PD 16-MAR-2001.  
 XX  
 XX 09-SEP-1999; 99FR-0011284.  
 XX  
 XX 09-SEP-1999; 99FR-0011284.  
 XX  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Beck A, Klinguer C, Nguyen TN;  
 XX  
 XX WPI; 2001-267782/28.  
 XX  
 DR N-PSDB; AAF84711.  
 XX  
 XX Use of quaternary aliphatic ammonium salt and immunogen or antigen to  
 PT combat respiratory syncytial virus infections -  
 XX  
 XX Claim 16; Page 22-23; 35pp; French.  
 XX  
 CC The specification describes a method which uses quaternary aliphatic  
 CC ammonium salts together with an immunogen or antigen to treat syncytial  
 CC virus infections. The combination of the salt with the antigen or  
 CC immunogen improves immunogenicity and equilibrates the Th1/Th2 immune  
 CC response. The method is used for the treatment of respiratory syncytial  
 CC virus infections. The present sequence represents a fusion protein,  
 CC comprising an albumin binding domain of the G protein of Streptococcus  
 CC sp. fused to a G protein fragment of respiratory syncytial virus (VRS,  
 CC long version). The fusion protein is used as an antigen in the method  
 CC of the invention.  
 XX  
 XX Sequence 349 AA;  
 SQ

Query Match 45.7%; Score 562; DB 22; Length 349;  
 Best Local Similarity 78.1%; Pred. No. 2.1e-36;  
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

QY 37 NLSEITSQTT--ILASTT-----PGVKSNOPTTVTKTKNTTTTQTPSKPTTKQKQKP 89  
 Db 216 NGKTLGKETTTEAADAATARSFNFPILNSM---TVTKTKNTTTTQTPSKPTTKQKQKP 272  
 QY 90 PNKPNDHFHEVFNFPVPCISCSNNPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 149  
 Db 273 PNKPNDHFHEVFNFPVPCISCSNNPTCWAICKRIPNKKPGKTKTKTKPTTKTKKDL 332  
 QY 150 KPQTTKPKVEPTTKPTE 166  
 Db 333 KPQTTKPKVEPTTKPVD 349

RESULT 15  
 AAR8253  
 ID AAR8253 standard; peptide; 101 AA.  
 XX  
 AC AAR8253;  
 XX  
 DT 15-MAY-1996 (first entry)  
 XX  
 DE RSV subgroup A wild type residues 130-230 (G2A clone).  
 XX  
 KW Immunogenic peptide; protein G; respiratory syncytial virus; adjuvant;  
 KW carrier protein; membrane lipopolysaccharide; LPS; Klebsiella pneumoniae;  
 KW divalent cation; detergent; anion-exchange chromatography; infection.  
 XX  
 OS Respiratory syncytial virus.  
 XX  
 XX FR2718452-A1.  
 XX  
 PD 13-OCT-1995.  
 XX  
 XX 06-APR-1994; 94FR-0004009.  
 XX  
 XX 06-APR-1994; 94FR-0004009.  
 XX  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PI Baussant T, Binz H, Thien NN, Trudel M;  
 XX  
 DR WPI; 1995-353189/46.  
 DR N-PSDB; AAT03486.  
 XX  
 XX New respiratory syncytial virus polypeptide(s) for vaccine prodn. -  
 PT esp. by conjugation with new klebsiella pneumoniae p40 protein  
 XX  
 XX Disclosure; Page 23; 38pp; French.  
 XX  
 CC The sequence of the wild type amino acid residues 130-230 from the  
 CC protein G of the respiratory syncytial virus (RSV) subgroups A. The  
 CC sequence was used to synthesise the immunogenic peptides AAR8245-52  
 CC based on residues 174-187 of the subgroup A and B RSVs. The peptides are  
 CC pref. conjugated to a novel carrier protein (the p40 protein; see  
 CC AAR8257) derived from a membrane lipopolysaccharide (LPS) fraction from  
 CC Klebsiella pneumoniae. The LPS fraction was isolated by precipitating the  
 CC Klebsiella membrane LPSs with a divalent cation and detergent,  
 CC subjecting the recovered proteins to anion-exchange chromatography to  
 CC obtain an immunological adjuvant and linking the p40 protein to the above  
 CC peptides. The conjugates are useful in the treatment of RSV A or B  
 CC infections.  
 XX  
 XX Sequence 101 AA;  
 SQ

Query Match 45.4%; Score 558; DB 16; Length 101;  
 Best Local Similarity 99.0%; Pred. No. 1e-36;  
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 64 TVTKTKNTTTTQTPSKPTTKQKQKPKNKPNNDHFHEVFNFPVPCISCSNNPTCWAICKRI 123  
 Db 1 TVTKTKNTTTTQTPSKPTTKQKQKPKNKPNNDHFHEVFNFPVPCISCSNNPTCWAICKRI 60  
 QY 124 PNKKPGKTKTKTKTKTKTKDLKPQTTKPKVEPTTKP 164

Db 61 PNXKPGKTTTKPTKRPFTFKTKRKHKPOFTKPKVEPTTKP 101

Search completed: October 29, 2003, 17:38:01  
Job time : 53.966 secs